

Univerzita Karlova
Přírodovědecká fakulta

Studijní program: Chemie
Studijní obor: Biofyzikální chemie



Bc. Josef Šulc

Příprava a charakterizace vazebných proteinů mimikujících epitopy
protilátek neutralizujících virus HIV-1

Preparation and Characterization of Protein Binders Mimicking Epitopes of
HIV-1 Neutralizing Antibodies

Diplomová práce

Školitel: RNDr. Petr Malý, CSc.

Praha, 2021

Prohlášení:

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením školitele RNDr. Petra Malého, CSc. a konzultanta Mgr. Milana Kuchaře, Ph.D. a všechny použité prameny jsem řádně citoval. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 15.8.2021

Podpis:

Poděkování

Za obrovskou pomoc při vypracování experimentální části své práce a za dlouholetou pomoc a vědecké vedení děkuji Mgr. Milanu Kuchařovi, Ph.D. a RNDr. Petrovi Malému, CSc. Za pomoc s experimenty, inspiraci a přínosné diskuze děkuji Ing. Haně Petrokové, Ph.D. a Mgr. Veronice Daniel Liškové, Ph.D. Stejně tak patří mé velké díky všem ostatním členům Laboratoře inženýrství vazebných proteinů, Tatsianě Charnavets, Ph.D. z Centra molekulární struktury a celému osazenstvu Biotechnologického ústavu AV ČR.

Abstrakt

Tři dekády trvající pandemie viru HIV stála život již desítky milionu lidí, další desítky milionů se s touto nevléčitelnou chorobou potýkají do dnešního dne. Současné neúspěchy ve vývoji ochranné vakcíny jsou způsobeny zejména extrémní mutační schopností viru, jeho velmi efektivním molekulárním štítem odrážejícím útoky imunitního systému a jeho nesmírnou variabilitou. Významným přelomem posledních let byl popis řady tzv. široce neutralizujících protilátek proti HIV-1, které nesou velmi účinnou a širokou neutralizační odezvu. Dosud se však nepodařilo nalézt cestu k navození vysoké produkce podobně účinných neutralizačních protilátek, které by ochránili imunizované jedince před infekcí a případným rozvojem syndromu AIDS. Tato práce popisuje výsledky výzkumu, který vedl k *in vivo* i *in vitro* prokázání funkčnosti konceptu tzv. proteinové mimetiky, tedy schopnosti kopírovat epitop virových povrchových proteinů, a tak vybudit účinnou imunitní odpověď ve formě produkce cílených široce neutralizujících protilátek. Účinek v tomto přístupu přináší rekombinantní vazebné proteiny na bázi proteinového scaffoldu Myomedin. Tato práce popisuje výběr a charakterizaci vazebných proteinů mimikujících epitopy nejúčinnější široce neutralizující protilátky 10E8. Ukazuje, že vazebné afinity vybraných Myomedinových variant jsou v řádech desítek nM (MLA158 $K_d = 10,8$ nM), že vazebná afinita k isotypové kontrolní protilátce je řádově nižší a že existuje výrazný rozdíl mezi vazbou vybraných variant a nemutovanou formou mateřského proteinu myomesin-1. Práce se zabývá selekcí a expresí těchto proteinů, purifikací a základními biofyzikálními charakterizacemi. Myší séra po vakcinaci zde vybranými a charakterizovanými proteinovými variantami prokázala schopnost neutralizace většiny pseudovirů HIV-1 z reprezentativního panelu. *In silico* výpočetní část se snaží objasnit strukturu vazebných proteinů, charakter jejich vazeb na 10E8 protilátku a poskytuje popis konkrétních interakcí, které mohou v biologických systémech přinášet kýženou imunitní odpověď. Tyto výsledky mají potenciál přinést další vylepšení mimikujících vazebných proteinů, a tak přiblížit vývoj vakcíny proti HIV-1 o krok blíže realitě.

Abstract

For three decades, the ongoing HIV pandemic has taken the lives of tens of millions of people. Still, more tens of millions are fighting this incurable disease today. Current failures in combating this global problem are caused mainly by the virus's extreme ability of mutation, its very effective molecular shield which repels the immune system's attacks, and its immense variability. A breakthrough, achieved relatively recently, is the discovery of the so-called broadly neutralizing antibodies against HIV-1, which carry a very efficient and broad neutralizing response. So far, it's not known how to elucidate the production of these antibodies in the infected hosts to quell or altogether eliminate the virus. This work deals with experimental results, which led to both *in vivo* and *in vitro* proof-of-concept of the so-called protein mimetics, the ability to imitate viral surface epitopes, and therefore stimulate an efficient immune response carried by targeted broadly neutralizing antibodies. This effect is mediated by recombinant binding proteins, based on the Myomedin scaffold. This work describes the selection and characterization of these binding proteins mimicking the epitopes of one of the most effective broadly neutralizing antibodies, 10E8. It shows that the binding affinities of selected Myomedin proteins are on the order of tens of nM (MLA158 $K_d = 10,8$ nM), that the affinity towards isotype control antibody is orders of magnitude less, and that there's a significant difference in binding between these selected variants and the paternal wild-type myomesin-1 protein. This work describes the selection and expression of these proteins, their purification and basic biophysical characterizations. Blood sera of mice, vaccinated by protein variants selected and characterized in this work, were able to effectively neutralize most HIV-1 pseudoviruses selected from a representative set. The *in silico* computational part of this work tries to explain the structures of the binding proteins, their interactions with the 10E8 antibody, and describes specific interactions which could be the cause of the observed biological effect. Based on knowledge brought by these results, mimicking binders could be further improved, bringing the development of an HIV-1 vaccine one step closer to reality.

(in Czech)

Klíčová slova

HIV, protilátka 10E8, vakcína, proteinový scaffold, kombinatoriální knihovna, vazebný protein, homologní modelování, dokování, MM/PBSA, umbrella sampling, mapy interakčních energií

Keywords

HIV, 10E8 antibody, vaccine, protein scaffold, combinatorial library, binding protein, homology modelling, docking, MM/PBSA, umbrella sampling, interaction energy maps

Seznam zkratek

bnAbs	Široce neutralizující protilátky proti HI <i>Broadly neutralizing antibodies</i>
BSA	Hovězí sérový albumin <i>Bovine serum albumin</i>
CD	Cirkulární dichroismus <i>Circular dichroism</i>
CG	<i>Coarse-grained</i>
COM	Těžiště molekuly <i>Center of mass</i>
DLS	Dynamický rozptyl světla <i>Dynamic light scattering</i>
DSF	Diferenční skenovací fluorimetrie <i>Differential scanning fluorimetry</i>
ELISA	<i>Enzyme-linked ImmunoSorbent Assay</i>
Fab	Vázající fragment protilátky <i>Fragment antigen-binding</i>
HAART	Vysoce aktivní antiretroviróvá terapie <i>Highly active antiretroviral therapy</i>
HC	Hydrofobní štěrbina (10E8 protilátky) <i>Hydrophobic cleft</i>
HIV	virus lidské imunodeficiency <i>Human immunodeficiency virus</i>
HPLC	Vysokoúčinná kapalná chromatografie <i>High-performance liquid chromatography</i>
IgG	Imunoglobulin G <i>Immunoglobulin G</i>
MD	Molekulová dynamika <i>Molecular dynamics</i>
MM/PBSA	Molekulová mechanika spjatá s Poisson-Boltzmannovou rovnicí a povrchem (molekuly) <i>Molecular mechanics/Poisson-Boltzmann Surface Area</i>

Ni-NTA	Nikl v komplexu s kyselinou nitrilotrioctovou (na agarose) <i>Nickel nitrilotriacetic acid</i>
NNRTIs	Nenukleosidové inhibitory reverzní transkriptasy <i>Non-nucleoside reverse transcriptase inhibitors</i>
NRTIs	Nukleosidové/nukleotidové inhibitory reverzní transkriptasy <i>Nucleoside/nucleotide reverse-transcriptase inhibitors</i>
OD ₆₀₀	Optická hustota při vlnové délce 600 nm
PBC	<i>Periodic boundary conditions</i>
SASA	Solventu dostupný povrch proteinu (v kontextu MM/PBSA z něj vycházející algoritmus) <i>Solvent accessible surface area</i>
SEC	Gelová permeační chromatografie <i>Size-exclusion chromatography</i>
SIV	Virus opičí imunodeficience <i>Simian immunodeficiency virus</i>
US	<i>Umbrella Sampling</i>

Obsah

1	Úvod	3
2	Přehled literatury	6
2.1	HIV-1	6
2.1.1	Viry lidské imunodeficiencie	6
2.1.2	Vývoj léčby HIV-1 od minulosti po budoucnost.....	7
2.1.3	Molekulární pohled na HIV-1 – genom, proteom a struktura	9
2.1.4	Genetická diverzita HIV-1	13
2.1.5	Env protein a mechanismus úniku před imunitní odpovědí	14
2.1.6	Široce neutralizující protilátky bnAbs	18
2.2	Mimikující proteiny	21
2.2.1	Malé proteiny a jejich využití	21
2.2.2	Myomedin.....	23
3	Cíle práce.....	25
4	Materiál a metody.....	26
4.1	<i>In vitro</i> Experimentální část	26
4.1.1	Použité přístroje a materiál	26
4.1.2	Postup práce	30
4.2	<i>In silico</i> část.....	41
4.2.1	Proteinové sekvence	41
4.2.2	Použitý software, skripty a balíky.....	42
4.2.3	Modelování, selekce a simulace	43
5	Výsledky a diskuze.....	51
5.1	Selekce proteinových variant vázících 10E8 IgG a jejich charakterizace	51
5.1.1	Úvod.....	51
5.1.2	Charakter experimentální práce	52
5.1.3	ELISA, plošná selekce z proteinové knihovny.....	53
5.1.4	Záměna C-terminálního Avi-tagu za V5-tag	55
5.1.5	Produkce, purifikace proteinů pomocí Ni-NTA agarosy pro analýzu vazebné afinity.....	56
5.1.6	Analýza vazebné afinity metodou ELISA	58
5.1.7	Purifikace SEC.....	60

5.1.8	ELISA jednotlivých separovaných SEC frakcí	67
5.1.9	Biofyzikální metody a charakter myomedinových forem	68
5.2	Charakterizace proteinových variant metodami <i>in silico</i>	78
5.2.1	Celkový cíl.....	78
5.2.2	Homologní modelování	79
5.2.3	Vylepšení a validace modelů (MD simulace).....	80
5.2.4	Dokování a prvotní výběr modelů	86
5.2.5	MM/PBSA a MM/GBSA analýza	90
5.2.6	Coarse-grained Umbrella Sampling.....	97
5.2.7	Umbrella Sampling s atomistickým rozlišením.....	99
5.2.8	Analýza intermolekulových interakcí komplexů s MPER	100
5.2.9	Analýza intermolekulových interakcí komplexů s MLA proteiny	108
5.2.10	Souhrn analýz intermolekulových interakcí	121
5.2.11	Tetramer MLA proteinů.....	124
5.3	Souhrn a návrhy na vylepšení vazebných proteinů.....	127
6	Závěr.....	129
7	Seznam literatury.....	131
8	Příloha A – experimentální část	i
8.1	Analýza vazebné afinity metodou ELISA.....	i
8.2	ELISA jednotlivých SEC separovaných frakcí.....	iv
9	Příloha B – <i>in silico</i> část.....	vi
9.1	Orientace ve vizuálních podobách matic volných energií párových interakcí aminokyselin	vi
9.2	Materiál a metody	viii
9.3	Výsledky a diskuze	ix

1 Úvod

V současném světě zmítaném pandemií COVID-19 je více než kdy jindy pozornost veřejnosti upínána na vztah vědy a medicíny. Nové výzkumné směry jsou více než kdy jindy využívány jako obrana před zhoubnými patogeny, které i přes extenzivní vývoj posledních dekad stále sužují podstatnou část lidské populace. Jedním z takových patogenů je virus HIV-1, který i přes četné vědecké průlomy nebyl stále úspěšně vymýcen. Primární obranou linií viru je glykanový štít, chránící důležité povrchové proteiny před imunitní odezvou. Dalšími účinnými nástroji viru je extrémní variabilita a schopnost dlouhodobé latence s inkorporací virové genetické informace do hostitelského genomu. Naděje na konečné řešení tohoto rozsáhlého problému spočívají ve vývoji efektivní a dostupné vakcíny.

Možným řešením zmíněných nepříznivých vlastností viru je zcela nový přístup využívající tzv. mimikujících proteinů, postavených na stabilním, strukturu udržujícím jádře a na několika záměnných aminokyselinách. Tyto malé proteiny jsou vybírány z rozsáhlých proteinových knihoven na základě své schopnosti vázat jednu ze široce neutralizujících protilátek proti HIV-1, vzácných a velice účinných imunoglobulinů. Takovým výběrem jsou získány molekuly, které svým povrchem napodobují vlastnosti původních vysoce konzervovaných virových epitopů a nastavují lidskému imunitnímu systému terč, nechráněný glykanovým štítem ani vysokou variabilitou. Popsaným mechanismem vzniká imunitní odezva řádově silnější, než jakou by způsobil samotný virus.

V naší laboratoři jsme vyvinuli dvě takové knihovny vazebných proteinů založených na základu malých rozpustných proteinů či endogenních proteinových domén. První z knihoven bylo ABD (albumin binding domain). Tento peptid o délce 46 aminokyselin, z toho 11 záměn, poskytl několik kandidátních molekul s *in vivo* prokázanou účinností zmíněného konceptu mimikování. Druhou knihovnou, jíž se detailně zabývá tato práce, je Myomedin. Tento 111 aminokyselin dlouhý protein s 12 záměnnými pozicemi vykazuje vysokou teplotní stabilitu a díky svému původu v přirozeném lidském proteinu je pravděpodobně lidskými hostiteli velmi dobře tolerován.

Některé, v této práci vybrané a popsané, varianty Myomedinových proteinů cílících na nejlepší prozatím popsanou široce neutralizující protilátku 10E8 byly ověřeny pomocí *in vivo* experimentů. Ty prokázaly velmi dobré schopnosti myších sér neutralizovat reprezentativní panel variant viru HIV-1.

Práce se zabývá výběrem Myomedinových vazebných proteinů na základě vazebných afinit proti protilátce 10E8, jejich biochemickou, molekulárně-biologickou a biofyzikální charakterizací a v neposlední řadě jejich rekombinantní produkcí. Využity byly standardní techniky molekulární biologie a biochemie (rekombinantní exprese v *E. coli*, elektroforéza, ELISA, gelová permeační chromatografie) a biofyzikální-chemie (DLS, CD, DSF, MS). Výsledky experimentů odhalily velice příznivé nanomolární vazebné afinity, přítomnost oligomerních stavů proteinu a další.

Druhou částí práce je aplikace *in silico* výpočetních metod pro bližší popis vazebných módů proteinů na 10E8 protilátku. S využitím homologního modelování, dokování, molekulové dynamiky a některých termodynamických analýz (MM/PBSA, Umbrella Sampling) jsou znovu popsány vazby v komplexech Myomedin-10E8, tentokrát však atomistickým pohledem. Prozkoumána je i vyřešená krystalová struktura komplexu 10E8 s nativním ligandem MPER, kde byly validovány a potenciálně rozšířeny původní výsledky autorů struktury. Tato část práce končí popisem konkrétních párových interakcí aminokyselin (pomocí vlastní přehledné a informativní metody reprezentace meziproteinových interakcí), které mohou nést biologickou aktivitu a otevírá cestu pro budoucí racionální optimalizaci vybraných vazebných proteinů.

2 Přehled literatury

2.1 HIV-1

2.1.1 Viry lidské imunodeficience

Viry lidské imunodeficience (dále jen HIV) náleží do čeledi retrovirů – RNA virů se schopností integrovat do hostitelského genomu sekvence virové cDNA pomocí reverzní transkripce. Jsou příslušníky rodu lentivirů (z latinského *lente* – pomalu) způsobujících chronické nemoci savců charakterizované dlouhými inkubačními periodami.

Popsány byly dva druhy virů lidské imunodeficience, HIV-1 a HIV-2. Z mnoha důvodů je globální pozornost věnována především druhu HIV-1, zejména díky jeho výrazně vyšší virulenci, infekčnosti a globální prevalenci. HIV-2 je lokalizován převážně v regionu západní Afriky, vykazuje oproti HIV-1 sníženou schopnost přenosu a zřídka kdy dospěje do nemoci, projevující se jako syndrom získaného selhání imunity (AIDS), jelikož je většina infekcí asymptomatických.¹

Na konci roku 2018 bylo dle dostupných informací Světové zdravotnické organizace těmito dvěma viry nakaženo odhadem 37,9 milionu lidí, tj. o 1,7 milionu více, než předchozí rok. Za celou dobu sledování epidemie onemocnělo přibližně 74,9 milionu a nastalo 32,0 milionu úmrtí. V souvislosti s virem způsobeným syndromem AIDS je nejčastější proximální příčinou úmrtí tuberkulóza.²

Současné poznání popisuje vznik HIV jako sérii několika nezávislých zoonotických infekcí viry opičí imunodeficience (SIV). Dle genetické příbuznosti se viry HIV-1 běžně dělí na typy M, N, O a P. Typ M je majoritním pandemickým typem dále děleným na 9 subtypů: A-D, F-H, J a K. Z těchto subtypů mají největší zastoupení v lidské populaci subtypy C a B. Subtyp C je predominantním subtypem HIV-1 v regionech Afriky a Indie; subtyp B je dominantní v západní Evropě, Americe a v Austrálii. Kromě zmíněných subtypů se zároveň vyskytují cirkulující rekombinantní subtypy.³ Přímým předchůdcem pandemického typu HIV-1 M byl s největší pravděpodobností virus SIVcpzPtt napadající

šimpanze učenlivého (*Pan troglodytes*), k jehož přenosu na člověka došlo ve dvacátých letech dvacátého století v Kongu.^{4,5}

2.1.2 Vývoj léčby HIV-1 od minulosti po budoucnost

Závažnost celosvětové pandemie HIV-1 začala být patrná okolo roku 1983, kdy byl virus poprvé izolován.⁶ Prvním významným molekulárním cílem léčby HIV-1 byl pro virus klíčový enzym – reverzní transkriptasa. První generací léků cílících na tento enzym byly terminátory transkripce založené na analogách nukleosidů a nukleotidů (NRTIs) – prvním takovým léčivem byl 3'-azido-2'-deoxythymidin (ZDV), schválený ve Spojených Státech roce 1987.⁷ Klíčovou vlastností této sloučeniny a léčiv ze třídy NRTI obecně je selektivita na virovou reverzní transkriptasu v porovnání s endogenními polymerasami.⁸

Při diskusi o léčbě nákazy virem HIV musíme rozlišovat dvě různé situace. Jednou situací je léčba již rozvinutého AIDS, ve kterém dokázal ZDV posunout mediánovou dobu úmrtí z 9,6 měsíců u neléčených pacientů na 21,2 měsíců.⁹ Druhou situací je snaha o zabránění rozvinutí infekce HIV do smrtelného stavu AIDS, který je možný pouze při dostatečném potlačení replikace viru. ZDV, stejně jako jiná NRTI sama o sobě nejsou schopna uspokojivě tohoto cíle dosáhnout kvůli vysoké rychlosti mutací reverzní transkriptázy a celého viru. Rezistentní kmeny viru, nesoucí čtyři aminokyselinové záměny v genu pro reverzní transkriptasu, se během terapie ZDV mohou začít objevovat již po šesti měsících.¹⁰

Efektivita léčby a její toxikologické vlastnosti byly částečně zlepšeny s nástupem kombinační terapie, zahrnující podávání dvou různých NRTI najednou. Tato léčba se ukázala jako výrazné zlepšení původní léčby založené pouze na podávání ZDV.^{11,12} I přes tato vylepšení stále nebylo dosaženo dlouhodobé suprese viru. Infekce HIV byla stále rozsudkem smrti, byť odloženým.

Nová éra účinných terapií proti infekci HIV nastala až s objevem alternativních tříd léčiv. Cílem zásahu jedné z nich zůstala reverzní transkriptasa, která však místo kompetitivní inhibice analogem substrátu začala být cílena nekompetitivně látkami vážícími se do alosterického místa enzymu. Prvním léčivem z této třídy nenukleotidových inhibitorů

reverzní transkriptasy (NNRTIs) byl nevirapin schválený v roce 1996. Výsledky studie INCAS provedené v roce 1998 ukázaly, že NNRTI nevirapin podávaný v trojkombinaci s dvěma NRTI (v rámci studie se zidovudinem a didanosinem), způsobil po šesti měsících od zahájení léčby u 79 % pacientů nekvantifikovatelně nízké koncentrace virové RNA. Zároveň byla oproti jakékoliv dvojkombinaci NRTI výrazně omezená úmrtnost a celková koncentrace viru v krvi. Takto výrazné omezení virové replikace znamená i výrazné omezení jeho schopností mutace, což snižuje i možnost vzniku rezistence.¹³

Nejvýznamnější třídou nově syntetizovaných léčiv byly inhibitory proteasy HIV, z nichž první, saquinavir, byl schválen v roce 1995.⁷ Díky nepříznivé metabolizaci saquinaviru enzymem CYP3A4 z rodiny cytochromů P450 musel být kombinován s odlišným inhibitorem proteasy ritonavirem, který zároveň působí jako velmi účinný inhibitor zmíněného enzymu. Kombinace těchto dvou léčiv společně se dvěma léčivy ze třídy NRTI se ukázala jako velice efektivní.¹⁴ Saquinavir, společně s účinnějším indinavirem, schváleným v roce 1996 daly vzniknout novému průlomovému typu terapií HAART, vysoce aktivním antiretrovirovým terapiím.⁷ Studie zabývající se trojkombinací látek indinaviru, ZDV a lamivudinu ukázala, že právě tato kombinace dokázala u většiny pacientů úspěšně potlačit virémii HIV po dobu nejméně tří let.¹⁵ Právě terapie HAART výrazně snížily mortalitu a morbiditu lidí nakažených HIV. Jen mezi roky 1995 a 1997 klesla dle studie provedené na 1255 amerických pacientech během doby pozorování mortalita na 100 osob z 29,4 na 8,8. Důvodem tohoto poklesu bylo zejména zahrnutí inhibitorů proteasy do léčebných postupů.¹⁶

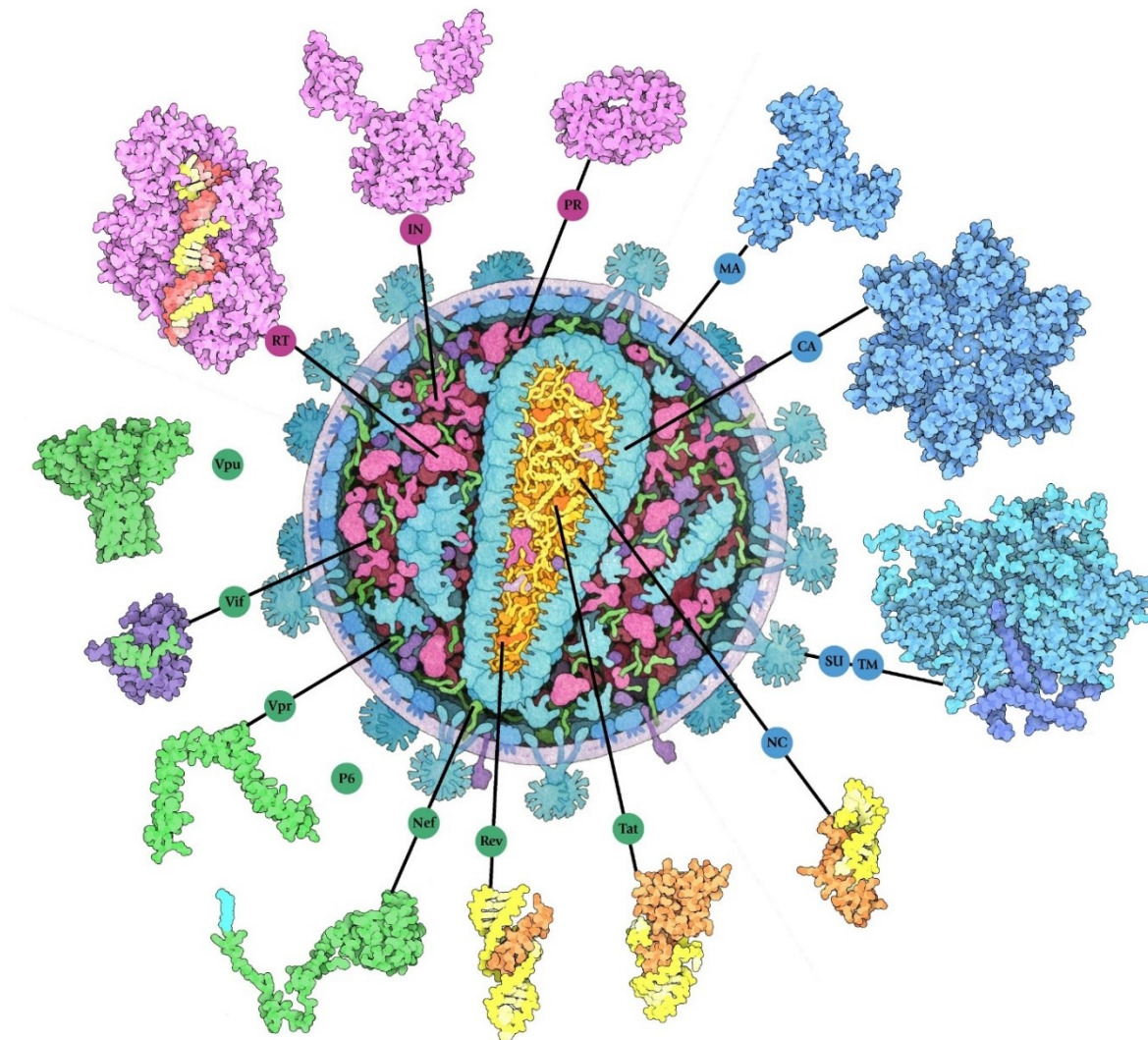
Prognóza nově nakažených pacientů je v současné době příznivá. Současná léčba zahrnuje obvykle kombinaci dvou NRTI s jedním lékem ze tříd NNRTI, inhibitorů proteasy nebo inhibitorů integrasy. Díky mnoha alternativám léčby, podávaným po vyvinutí rezistence na předchozí léčbu, s jednoduchým podáváním a oproti minulosti nízkou toxicitou se doba života pacientů s HIV blíží době života bez infekce.¹⁷ Snížení koncentrace viru v krvi zároveň znamená výrazný pokles nakažlivosti. HIV se zejména v rozvinutých zemích z letální nemoci stalo chronickým problémem a místo záchrany životů se do popředí dostávají i další problémy spojené s celosvětovou pandemií viru HIV. Jedním z těchto problémů je zejména enormní finanční zátěž spojená se zajištěním

léčby a diagnostických zákroků pro velké množství pacientů a zajištění léčby a diagnostiky v rozvojovém světě, kde je HIV ještě stále častou příčinou smrti.

Vývoj vakcíny proti HIV je nejlepším a možná jediným způsobem, jak definitivně vyřešit dekády trvající problém globální pandemie HIV. Staré přístupy vývoje vakcín se i přes četné snahy neukázaly jako účinné – zejména kvůli nezvyklé schopnosti viru mutovat, jeho integraci do hostitelské DNA a jeho nízké antigenicitě (tyto a další faktory jsou diskutovány v **kapitole 2.1.5**). Moderní imunoterapeutické přístupy mají však nesmírný potenciál poskytnout konečnou odpověď na vývoj nejen vakcíny proti HIV.

2.1.3 Molekulární pohled na HIV-1 – genom, proteom a struktura

Virion HIV-1 je sférického tvaru s průměrem přibližně 100 nm. Genom viru ve formě dvou identických kopií (+)-RNA řetězců je lokalizován uvnitř kuželovité nukleokapsidy tvořené proteinem CA (p24). Na každou molekulu RNA je vázána jedna molekula reverzní transkriptázy, molekula integrasy, molekula proteasy a nukleokapsidový protein NC (p7). Vnitřní stranu nosné struktury virionu tvoří vrstva molekul proteinu MA (p17), na které se nachází lipidová dvojvrstva tvořící obal virionu. Při pohledu na virion z vnější strany jsou patrné struktury do membrány zanořené trimerů glykoproteinu gp120. Transmembránovou kotvu, uchycující gp120 do membrány tvoří glykoprotein gp41.^{18,19} Schematický pohled na virion HIV-1 je zobrazen na **obrázku 1**.

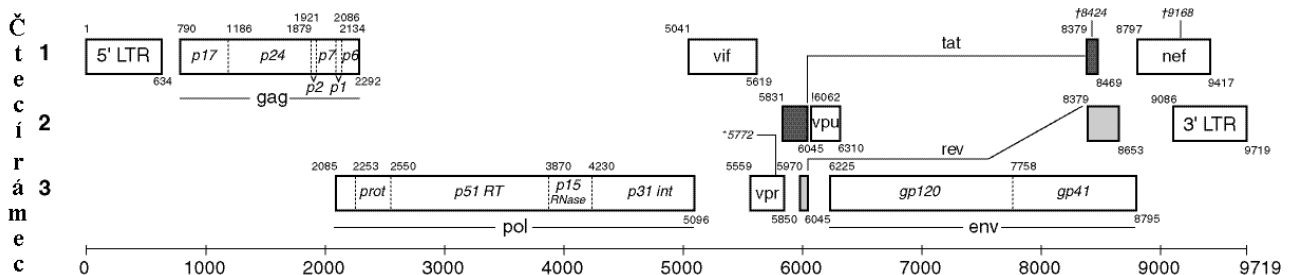


Obrázek 1 - Schematický nákres struktury virionu HIV-1 Fialově jsou znázorněny virových enzymy, modře strukturní proteiny, zeleně pomocné proteiny, žlutě krátké úseky jednořetězcové RNA a oranžově proteiny přímo interagující s RNA. Obrázek byl přejet z internetového zdroje a upraven autorem.²⁰

Genom viru je kódován ve formě dvou identických (+)-RNA řetězců, modifikovaných na 5'-konci čepičkou a na 3'-konci poly(A) úseky. V blízkosti 5'-konce je lokalizována (v pořadí od 5'-konce) repetice R, unikátní sekvence U5 a unikátní sekvence PB. Před 3'-koncem se nachází (v pořadí od 5'-konce) polypurinová sekvence PP, unikátní sekvence U3 a repetice R. Tato genomová RNA vzniká transkripcí provirové DNA pomocí RNA-polymerasy II hostitelské buňky.¹⁸

Provirová DNA vzniká v rámci životního cyklu viru reverzní transkripcí genomové RNA v hostitelské buňce, kam je následně integrována. Tato DNA má oproti genomové RNA odlišné konce, které jsou tvořeny sekvencemi LTR („long terminal repeats“).¹⁹ Na 5'-konce LTR se vyskytuje unikátní sekvence U3 následovaná repeticí R a unikátní sekvencí U5. Těmto elementům následuje přepis RNA sekvence od PB až do PP (směrem 5'-3'), následovaná opakováním celého LTR elementu (od 5'-U3, R, U5).¹⁸ 5'-konce LTR sekvence slouží jako promotor transkripce.¹⁹

Samotné exony obou forem genomu jsou tvořeny geny *gag*, *pol* a *env* kódujícími prekurzory proteinů viru. Důležitou regulační úlohu hrají doplňkové proteiny získané alternativním splicingem genů *pol* a *env*. V rámci genomu existují tři otevřené čtecí rámce („open reading frames“, ORF), generující dohromady, po štěpení proteasou, 15 virových proteinů. Elementy genomu jsou graficky znázorněny na **obrázku 2**. Jednotlivé proteiny kódované genomem HIV-1 jsou uvedeny v **tabulce 1**. Stručný popis jejich funkce je uveden v **tabulce 2**.



Obrázek 2 - Mapa elementů genomu HIV-1. Spodní osa zobrazuje páry bazí, obdélníky znázorňují proteinové produkty jednotlivých genů. V rámci HIV-1 existují tři možné čtecí rámce, které jsou vyjádřeny svíslou osou. Číslo v levém horním rohu obdélníků označuje první nukleotid START-kodonu, čísla ve spodním pravém rohu označují poslední nukleotid STOP-kodonu. Obrázek byl získán z internetového zdroje a upraven autorem.²¹

Tabulka 1 - Výčet genů kódovaných genomem HIV a jejich produktů. Genom obsahuje exony gag, pol a env. Alternativním splicingem oblasti překrývající pol a env exony je možno získat doplňkové proteiny, které jsou uvedeny v pravé části tabulky. V závorkách jsou uvedeny alternativní názvy proteinů vycházející z jejich molekulové hmotnosti v kDa. Poznámka – polyproteiny p160 jsou navzájem strukturně odlišné.

Gen	Prekurzorový polyprotein	Protein	Doplňkový protein
<i>gag</i>	Pr55gag (p55)	MA (p17)	Tat (p9/p14)
		CA (p24)	Rev (p14)
		NC (p7)	Vif
		LI (p6)	Vpu
<i>pol</i>	Pr160GagPol (p160)	RT + RNasa H	Vpr
		IN	Nef
		PR	
<i>env</i>	PrGp160 (p160)	SU (gp120)	
		TM (gp41)	

Tabulka 2 - Výčet proteinů exprimovaných virem HIV-1. Vyjmenováno je krátké shrnutí funkce proteinů.¹⁹

Protein	Popis
MA	Matrixový protein tvořící vnitřní vrstvu membrány
CA	Kapsidový protein tvořící kónickou kapsidu
NC	Nukleoprotein napomáhající RNA zaujmout stabilní strukturu. ²²
LI	Mediátor pučení viru z hostitelské buňky.
RT + RNasa H	Reverzní transkriptasa + RNasa degradující RNA z komplexu virové RNA/DNA získaného reverzní transkripcí
IN	Integrasa
PR	Proteasa štěpící Gag (p55) a Gag-Pol (Pr160GagPol) prekurzorové polyproteiny
SU	Membránový glykoprotein gp120
TM	Glykoprotein gp41 kotvící gp120 do membrány
Tat	Transaktivační protein
Rev	Regulátor sestřihu RNA
Nef	Negativně reguluje vystavení CD4 receptorů na povrchu napadené buňky aby nedošlo k infekci jinou částicí HIV-1.
Vif	Virový infekční faktor (Viral infectivity factor), způsobuje degradaci hostitelského enzymu APOBEC3G. ²³
Vpu	Virový protein unikátní (viral protein unique), indukuje degradaci CD4 receptoru
Vpr	Virový protein R, interaguje s proteinem LI a způsobuje infektivitu viru.

2.1.4 Genetická diverzita HIV-1

Genetická diverzita je výraznou vlastností typickou pro RNA viry, zejména pak pro viry HIV. Infekce tímto virem je charakterizována rychlou a chybovou replikací ústící ve velice diverzní virovou populaci schopnou odolávat selekčním tlakům. Tato variabilita je majoritní překážkou pro léčebné snahy jak antiretrovirové terapie, tak potenciální budoucí vakcíny. Zdrojů této diverzity je několik – vysoký virový turnover, vysoká schopnost mutace, možnost rekombinace a selekční tlak hostitelského imunitního systému.²⁴ *In vivo* byla zjištěna extrémní mutační rychlost $(4,1 \pm 1,7) \cdot 10^{-3}$ mutací v jedné bázi v rámci jednoho virionu. Jedná se o nejvyšší mutační rychlost ze všech biologických entit. Je však nutné poznamenat, že většina těchto mutací je pro virus letální.²⁵

Dva nejdůležitější zdroje jednotlivých mutací virového genomu jsou mutagenní faktory hostitelské buňky (APOBEC3G enzym katalyzující přeměnu deoxycytidinu na deoxyuridin ve virové (-)-DNA v rámci imunitní odezvy hostitele) a chybová reverzní transkriptasa viru (*in vitro* chybovost 1 z 6900 nukleotidů RNA templátu).^{26,27} Dalším zdrojem celkové diverzity HIV-1 je rekombinace usnadněná diploidii RNA genomu. Pokud je buňka nakažena více než jedním virem, může dojít v nově vzniklé virové partikuli k „zabalení“ dvou různých RNA genomů za vzniku heterozygotního viru. Tyto molekuly RNA mohou být reverzní transkriptasou přepsány do jedné chimérické virové DNA.²⁸

Jak bylo diskutováno v **kapitole 2.1.1**, HIV-1 je jedním ze dvou popsaných druhů virů lidské imunodeficience a dělí se na řadu typů a subtypů na základě genetické příbuznosti. HIV-1 se od HIV-2 liší přibližně v 54 % aminokyselin. Aminokyselinový mezi jednotlivými typy HIV-1 (M, N, O, P) je zhruba 41 %, rozdíl mezi jednotlivými subtypy (A-D, F-H, J, K) činí přibližně 18 %. V rámci 15 virových proteinů je mezi subtypy nejvíce konzervována aminokyselinová sekvence integrasy, nejvíce variabilní je naopak sekvence glykoproteinu gp120.²⁹

Z hlediska potenciální vakcíny proti HIV-1 je zajímavý poznatek, že významná variabilita v aminokyselinové sekvenci existuje zejména v pozicích sloužících jako epitopy pro CD4 T buňky a protilátky imunitního systému. Právě v těchto pozicích bylo v rámci subtypu B, rozšířeném zejména v zemích prvního světa, lokalizováno 919 (tj. 30 % z celku) variabilních pozic s diverzitou vyšší, než 12.9 %. Největší selekční tlak fixující mutace je zejména na proteiny mající interakce s lidskými proteiny. Těmi jsou obálkové proteiny gp120 a gp41 a regulační proteiny Tat a Rev. Naopak virové enzymy mají diverzitu nejmenší.²⁹ Tato variabilita a zároveň regionálně specifické rozšíření různých subtypů viru je výraznou překážkou pro vývoj univerzálně funkční vakcíny.

I přes vysokou diverzitu je možno najít konzervované sekvence (jak mezi typy/subtypy tak mezi jednotlivými viriony), zejména (ale ne exkluzivně) v proteinech CA, NC, PR, RT, IN, V_{rp} a v N-terminální doméně gp41. Na tyto konzervované sekvence proteinů PR, RT a IN cílí současná hojně užívaná léčiva, jak bylo diskutováno v **kapitole 2.1.1**. V současné době vyvíjeným alternativním přístupem jsou inhibitory proteinů kódovaných genem *gag* (zejména CA, NC) cílící obdobně na konzervované sekvence.³⁰

I přes zmíněnou diverzitu povrchových proteinů gp120 a gp41 obsahují tyto proteiny některé konzervované sekvence, které mohou sloužit jako terapeutický cíl.

2.1.5 Env protein a mechanismus úniku před imunitní odpovědí

Klíčovým proteinem HIV-1 pro vstup do hostitelské buňky skrze interakci s CD4 receptorem je povrchový glykoprotein gp120 a transmembránová kotva gp41, která fixuje gp120 na virové membráně. Kombinace těchto dvou glykoproteinů dohromady tvoří tzv. Env protein který existuje nativně v bohatě glykosylované trimerní formě, jejíž proteinová část sestává z trimeru gp120 a trimeru gp41. Glykany jsou převážně oligomanosového, díky sterickému bránění funkcí endoplasmatického retikula a manosidas nematurovaného typu a tvoří zhruba 50 % celkové váhy proteinu. Nejčastěji se objevují glykany typu Man₈₋₉GlcNAc₂.³¹ Mediánový počet připojených glykanů na povrchu gp120 je uváděn na 25.^{32,33} Glykosylace gp41 je od gp120 odlišná, její hustota je nižší, což zřejmě souvisí s rozdílným složením způsobeným rozdílnou maturací, kde menší hustota glykosylace poskytuje menší sterické bránění tohoto procesu. Glykosylace

gp41 se sestává primárně z endo- β -N-acetylglukosaminidase H resistantních komplexních typů glykanů.³⁴

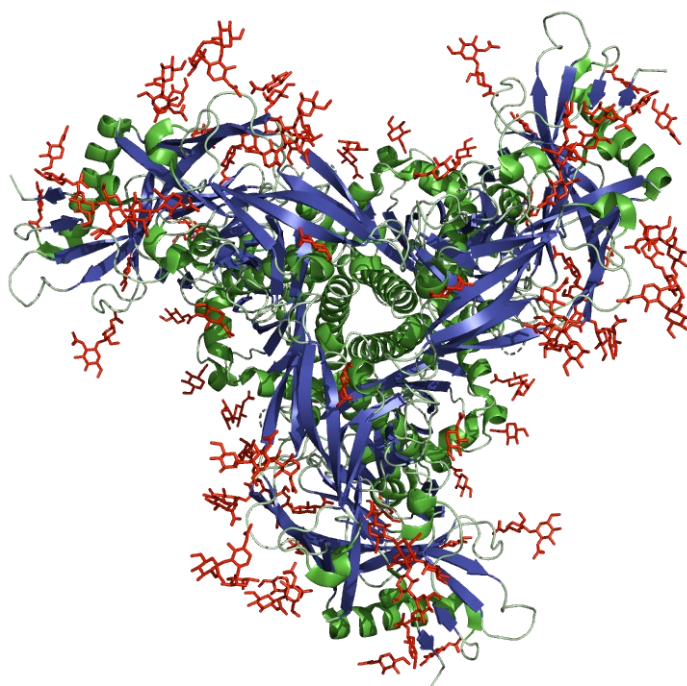
Nezvykle vysoká hustota glykanového štítu na gp120 omezuje maturaci během biosyntézy virionu, což zanechává region zvaný „intrinsic mannose patch“ složený z nematurovaných oligomanosových struktur. Glykosylace takového charakteru je mezi savčími glykoproteiny neobvyklá, což ji činí potenciálním cílem imunitní odpovědi. Tento region je primárním cílem tzv. široce neutralizujících protilátek bnAbs („broadly neutralizing antibodies“) v gp120 části Env.³⁵

Společně s diverzitou diskutovanou dříve je bohatá a variabilní N-glykosylace tohoto proteinu je jedním z hlavních mechanismů úniku HIV-1 před imunitní odezvou. Zmíněné glykany chrání relativně konzervované proteinové sekvence před protilátkovou odezvou, jsou nestabilní, variabilní a jsou samy o sobě velmi málo imunogenní. Cukerné epitopy mají inherentně slabou interakci s proteiny, existují navíc v rozdílných formách, což dále snižuje účinnost imunitní odezvy. Tato slabá interakce protilátek se sacharidovými řetězci je často zesílena multivalencí molekuly protilátky na glykan a část Env proteinu.³⁶

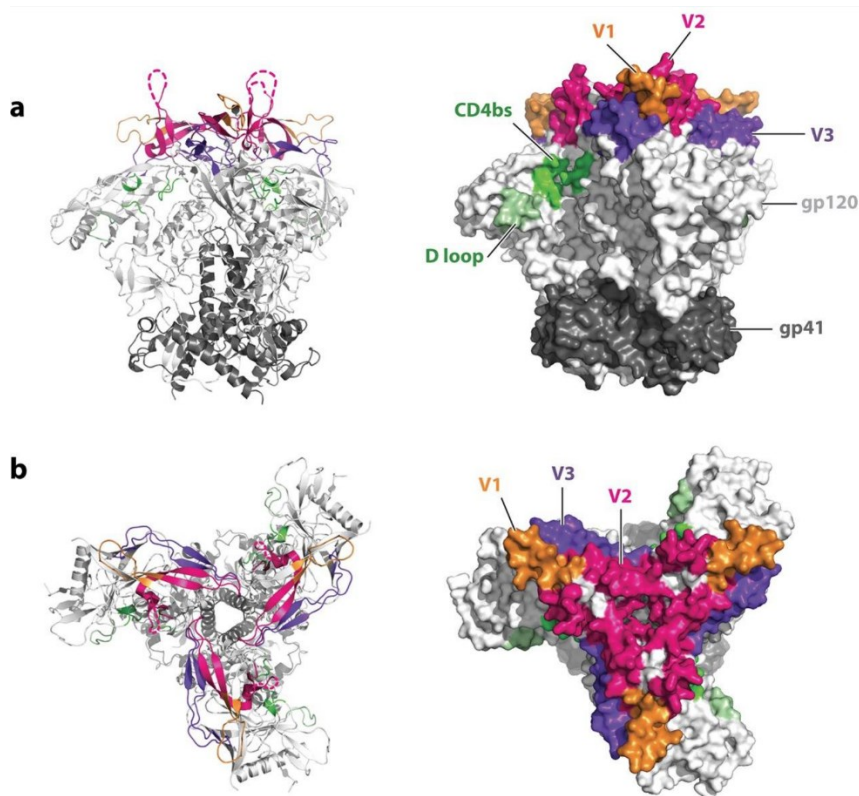
Glykany jsou navíc často dynamické a nacházejí se v rozdílných konformacích. Jelikož není glykosylace výrazně odlišná od hostitelské, mohou být tyto exogenní molekuly do jisté míry hostitelským imunitním systémem tolerovány a odezva proti nim může být vystavena negativní selekci díky možné auto-reaktivitě. Stále však existuje jejich odlišnost od endogenních glykosylů z hlediska jak složení, tak struktury (vysoké density), která může vyústit v určitou imunitní odpověď.

Neobvyklostí viru HIV-1 je nízký počet těchto „spike proteinů“, uváděný na 8-14 Env proteinů na virion.^{37,38} Nízká hustota těchto povrchových proteinů je jedním z faktorů negativně ovlivňujících imunitní odpověď proti viru, na druhou stranu umožňuje přístup protilátek ze všech stran glykoproteinu.³⁹

Podobu artificiálně připraveného konstruktu BG505.SOSIP664 (rekombinantní produkovatelná gp140 prekurzoru Env proteinu) zachycuje **obrázek 3**. Proteinovou část bez glykanů zobrazuje **obrázek 4**.



Obrázek 3 – BG505.SOSIP – rekombinantně exprimovatelná artificiální forma virového proteinu Env. Červeně jsou vyobrazeny glykany. (PDB: 4ZMJ)⁴⁰



Obrázek 4 – Struktura proteinové části Env trimeru HIV-1. Převzato.³⁹ a) pohled ze strany, b) pohled ze shora

Env protein s mohutným glykanovým štítem je jediným cílem široce neutralizujících protilátek proti HIV-1. Tyto protilátky jsou schopny neutralizovat vysoké množství různých HIV-1 variant a tvoří se u 10-30 % nakažených jedinců. Jejich séra pak mají různou potenci a šíři neutralizační schopnosti s ohledem na varianty HIV-1.⁴¹

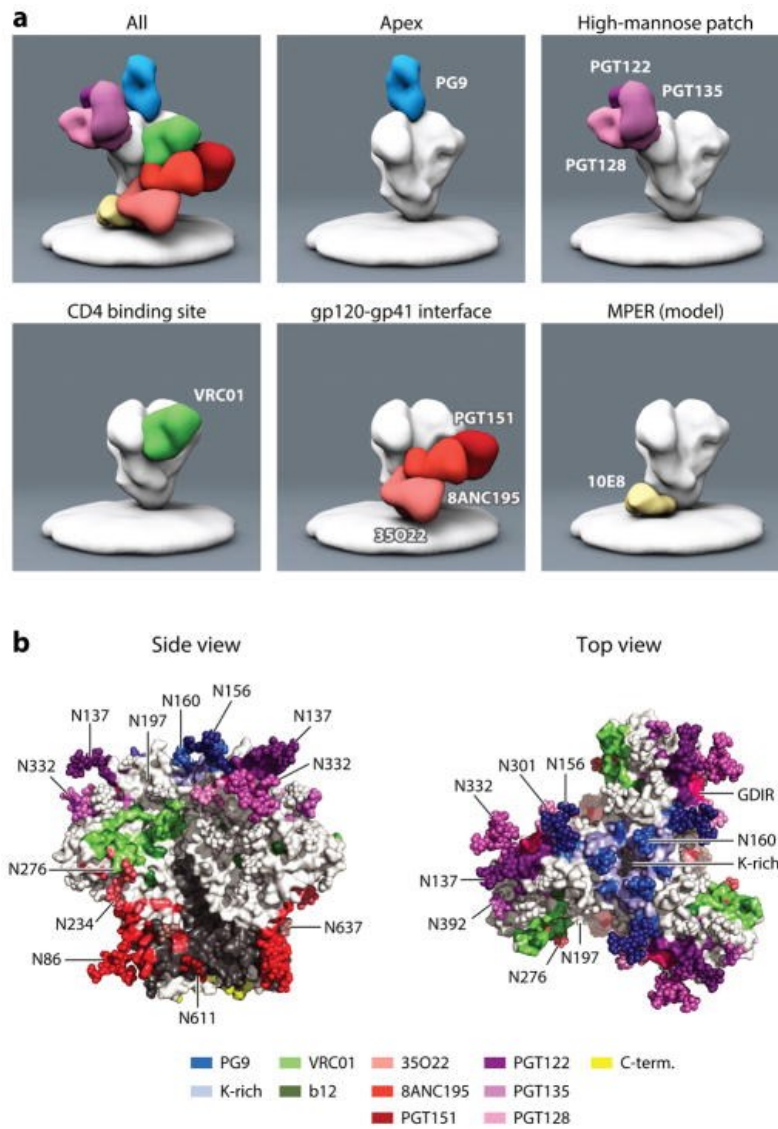
Kromě zmíněných na glykany cílících protilátek existuje i „hybridní“ typ protilátek cílících na gp120 epitopy V3 smyčky, a to protilátky vážící protein-glykanové epitopy rozpoznávané bnAbs PGT121 a PGT126, které byly v naší laboratoři také vybrány jako molekulární cíle pro selekci mimikujících Myomedinů. Společně s protilátkami specifickými pro epitopy CD4-vazebného místa na gp120 a pro MPER proteinový region gp41 patří tyto protilátky k nejúčinnějším z hlediska neutralizačního potenciálu. Důležitým faktorem pro vývoj vakcín založených na proteinech mimikujících epitopy gp120 glykanů je však diverzita gp120 a inserce/delece glykanů jako reakce na selekční tlak imunitního systému.³⁵

2.1.6 Široce neutralizující protilátky bnAbs

Po infekci virem HIV-1 se po 2-4 týdnech vyvíjí neúčinné, ne-neutralizující volné protilátky proti gp120 a gp41.⁴² Neutralizující protilátky se objevují v řádu měsíců a cílí nejčastěji na variabilní regiony gp120 (V1/V2, V3 loop) a jsou tedy specifické pro danou variantu viru. Toto umožňuje viru jednoduše porazit tento selekční tlak skrze mutace.^{43,44} Dle kritérií pro zařazení protilátky do kategorie bnAbs mohou tyto protilátky vzniknout v 10-50 % infekcí.^{45,46} Skutečně účinná odpověď skrze nejlepší bnAbs se však vyvíjí v řádu jednotek procent infikovaných („elite neutralizers“).⁴⁷ Právě indukce takových účinných a širokospektrých protilátek je cílem snah o vývoj preventivní vakcíny proti HIV-1.

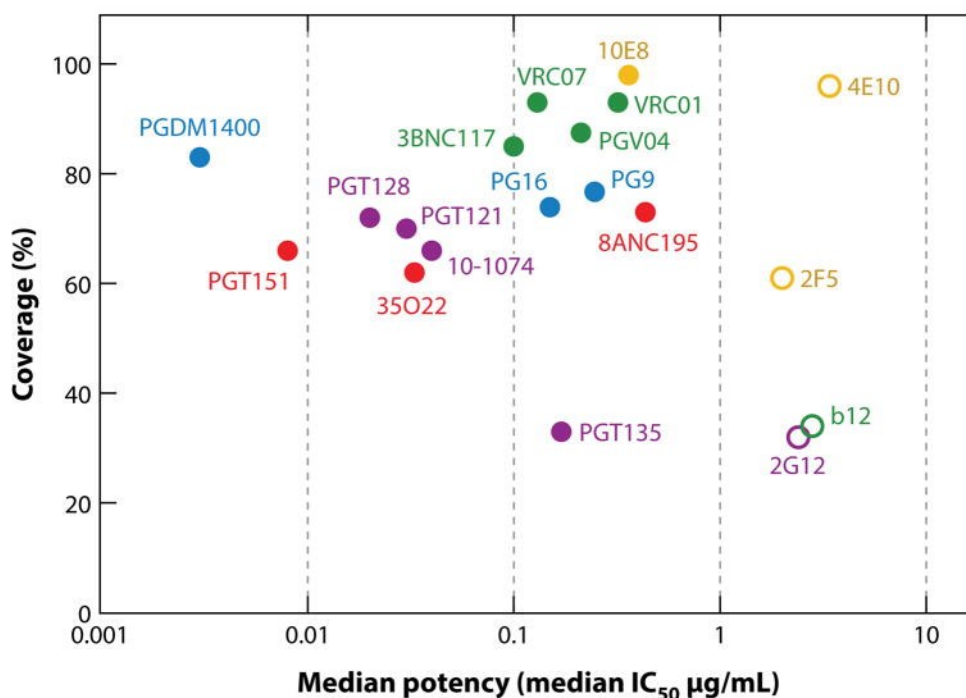
Dosud popsané široce neutralizující protilátky proti HIV-1 mohou mít za cíl několik částí povrchového Env proteinu. Vizualizace vazebných míst je zobrazena na **obrázku 5**.

- a) CD4-vazebné místo (protilátky b12, **VRC01**, PGV04)
- b) MPER region („membrane proximal external region“) gp41 (**10E8**, 4E10, 2F5)
- c) Glykanový štít (N332 glykan a V3 loop - **PGT121**, PGT124, **PGT126**, PGT128, PGT135 ; N160 glykan a V1/V2 loop – PG9, PGT145 ; glykany rozhraní gp41-gp120 – PGT151) – typicky interakce jak s glykany, tak s protein. Tyto protilátky jsou typicky neobvyklé přítomností dlouhých CDRH3 regionů sloužících k průniku glykanovým štítem k dosažení vazby s Env proteinem.³⁴



Obrázek 5 – Vazebná místa Env proteinu (gp41+gp120) HIV-1 pro široce neutralizující protilátky. Převzato.³⁹

Přehled bnAbs z hlediska pokrytí variant HIV-1 a neutralizační efektivity je zobrazen na obrázku 6.



Obrázek 6 – Přehled široce neutralizujících protilátek proti HIV-1 z pohledu pokrytí variant viru (osa Y) a mediánové IC₅₀ (osa X). Převzato.³⁹

Speciální pozornost je v této práci věnována protilátce 10E8, která byla využívána v praktické části této práce. Neutralizační šíře (procento virů neutralizovaných alespoň z 80 % při maximální koncentraci protilátky $50 \mu\text{g} \cdot \text{ml}^{-1}$) je u ní nejvyšší ze všech dosud popsaných protilátek – 97 %. Její hodnota $\text{IC}_{80} = 2,05 \mu\text{g} \cdot \text{ml}^{-1}$ je zároveň nejnižší mezi dosud popsanými protilátkami cílenými na MPER epitop. Společně s absencí autoreaktivity této protilátky poskytuje její schopnost neutralizovat většinu typů viru a nízká účinná koncentrace ideální cíl pro snahy o vývoj vakcíny proti HIV-1.⁴⁸

Tato protilátka specifická pro epitop MPER oblasti gp41 má slabou afinitu k nativnímu trimeru viru a předpokládá se, že neutralizuje skrz vazbu fúzního intermediátu zformovaného po vazbě viru na CD4 receptor obdobně jako protilátky 2F5 a 4E10 cílící také na MPER.^{49,50} Důkazem pro tuto vazebnou závislost je i pozorovaná zesílená vazba Env proteinu se séry po imunizaci myši variantami MLA po introdukci volného CD4 receptoru.⁵¹

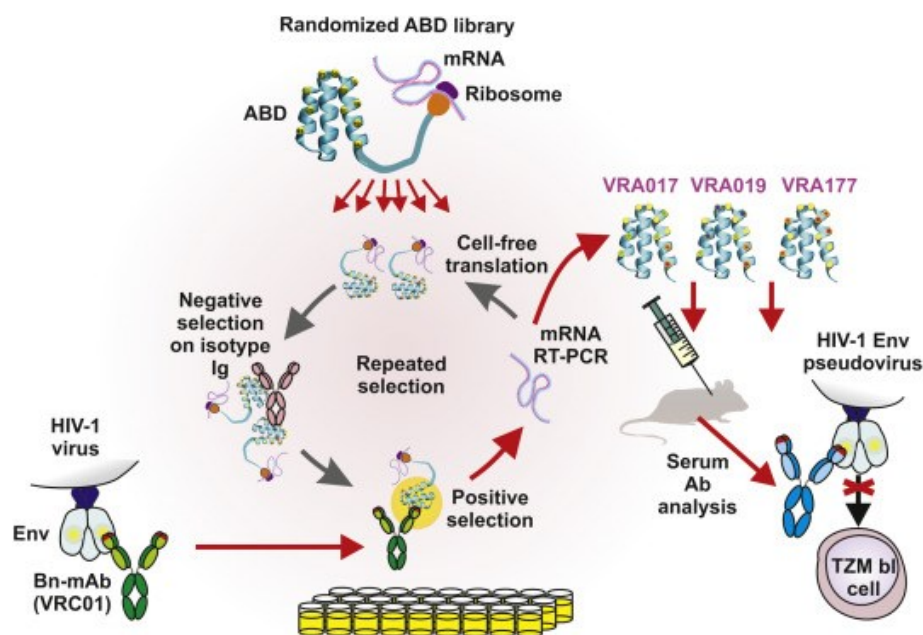
2.2 Mimikující proteiny

2.2.1 Malé proteiny a jejich využití

Jak bylo diskutováno dříve, v současné době nejnadějnější strategií boje proti HIV-1 je indukce imunitní odpovědi skrze široce neutralizující protilátky, které jsou tvořeny tzv. „elite neutralizers“, vzácnou podmnožinou pacientů schopných tvořit efektivní protilátky schopné neutralizovat široké spektrum možných typů viru. Klíčovým krokem pro umělou indukci a zesílení imunitní odezvy tohoto typu je vytvoření vakcíny, která by obešla inherentní nízkou imunogenicitu virových epitopů a vybudila imunitní odpověď. Možnou odpovědí na otázky spojené s vývojem by byla nově navržená strategie imunizace proteinovými scaffoldy mimikujícími povrch virových epitopů.

Proteinové scaffoldy jsou odvozeny od stabilních, typicky lidských či bakteriálních proteinů. Mají konstantní regiony, tedy konzervované sekvence aminokyselin tvořící stabilní terciární strukturu a vytyčené variabilní pozice, ve kterých může docházet k aminokyselinovým záměnám bez změny terciární struktury a ideálně stability. Díky těmto záměnným aminokyselinám je možné vygenerovat kombinatoriální knihovny obsahující velké množství variant, které mohou být selektovány pro své vazebné schopnosti na určený cíl. Cílem této snahy je mj. získat specificky vázající, afinní, stabilní a *in vivo* dobře tolerované proteiny, které mohou být snadno produkovány rekombinantními metodami. Díky širokému potenciálu využití dnes existuje scaffoldů celá řada, významné jsou scaffoldy Affibody, Anticalin, DARPin a mnohé další.⁵²

Celý koncept konkrétní aplikace proteinových scaffoldů pro vývoj vakcíny proti HIV-1 je ve zkratce popsatečný, pro konkrétní případ scaffoldového proteinu ABD, na **obrázku 7**.



Obrázek 7 – Schéma tvorby a výběru epitopy mimikujících proteinů založených na scaffoldu ABD. Převzato.⁵³

Je vytvořena komplexní proteinová knihovna možných vazebných proteinů založených na proteinovém scaffoldu s několika proměnnými aminokyselinami a stabilním aminokyselinovým „jádre“ udržujícím terciární strukturu a stabilitu. Jako cíl pro selekci (např. metodou ribosomálního displeje) je využito imobilizované monoklonální protilátky. Pro negativní selekci je použito směsi protilátek stejného isotypu. Po několika kolech selekce je uvolněna mRNA, reverzní transkripcí přepsána na DNA a vložena do plasmidů. Na základě takto získaných proteinů, schopných vázat cílovou protilátku, jsou provedeny další selekce na základě poměrů vazebných afinit k cílové a isotypové protilátce. Vybrané varianty jsou charakterizovány z hlediska detailněji stanovených vazebných afinit. Mohou být dále upravovány pro zvýšení stability, například introdukcí náhodných bodových mutací či cílenou záměnou aminokyselin, a dále analyzovány. Nejlepší výsledné varianty jsou testovány *in vivo* z hlediska indukce protilátek neutralizujících pseudoviry HIV-1. Tímto krokem proběhne zároveň další selekce, která již vybere proteiny s prokázanou biologickou aktivitou.

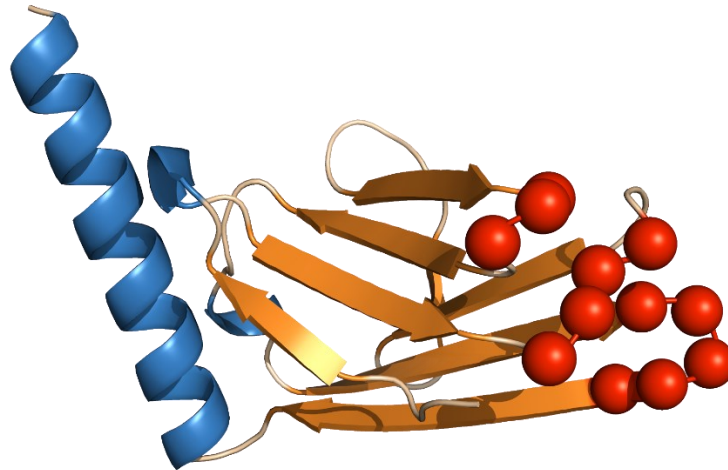
Jako důkaz funkčnosti zmíněného konceptu slouží několik *in vivo* experimentů provedených v naší a spolupracujících laboratořích. Pro vazebné proteiny založené na ABD scaffoldu (VRA017, VRA019, VRA177) mimikující epitop protilátky VRC01 bNAb bylo prokázáno, že myši po aplikaci VRA proteinů produkují virus neutralizující séra, která byla testována na panelu pseudovirů HIV-1. Zmíněné vazebné proteiny zároveň v *in vitro* experimentech prokazují dobrou vazebnou afinitu a specifitu, dobrou stabilitu a stabilní terciární strukturu. Tři imunizace myši ABD vazebnými proteiny VRA017, a VRA177 vedly prokazatelně k neutralizační schopnosti myších sér proti vybraným pseudovirům a k vývoji protilátek vázajících gp120. Toto přineslo důkazy pro možnou funkčnost výše popsaného kontextu využití mimikujících proteinů k vybuzení imunitní odpovědi proti HIV-1.⁵³

2.2.2 Myomedin

Proteinový scaffold Myomedin vychází z domény 10 lidského svalového proteinu myomesin-1. Tato doména byla využita k vyvinutí malého vazebného proteinu zvaného Myomedin pro generování knihoven vazebných variant s aminokyselinovými záměnami na smyčkách představující nestrukturované úseky mezi β -listy.

Tento malý vazebný protein má 111 aminokyselin, je bohatý na β -listy a má jeden terminální α -helix. Pozice aminokyselinových záměn byly vybrány na základě *in silico* screeningu mutací se sledováním změn volné energie po mutaci aminokyselin. Tato knihovna je v současné době využívána pro generování vazebných proteinů pro vazbu do variabilních oblastí široce neutralizujících protilátek proti HIV-1, PGT121, PGT126, 10E8.

Cílem bylo najít ty vazebné varianty, které budou mimikovat epitopy těchto neutralizujících protilátek. Struktura Myomedinu je zobrazena na **obrázku 8**.



Obrázek 8 – Struktura proteinového scaffoldu Myomedin. Oranžově jsou vyznačeny β – listy, modře α – helix, hnědě nestrukturované úseky a ohyby. Aminokyselinové záměny jsou označeny červenými koulemi.

Myomedin je vhodný pro rychlou přípravu účinných vazebných proteinů pro diagnostické i terapeutické účely. Předpokládá se, že vzhledem k endogennímu charakteru proteinu by měl být lidmi dobře tolerován (sám o sobě by ideálně neměl být imunogenní). Teoretická komplexita knihovny připravené na základě aminokyselinových záměn je $2 \cdot 10^{15}$ variant.⁵¹

3 Cíle práce

Cílem *in vitro* experimentální části této práce bylo nejprve separovat, produkovat a vybrat proteiny kódované plasmidovými DNA získanými selekcí z kombinatoriální knihovny, na základě rozdílu vazebné afinity k 10E8 protilátce a isotypové kontrole IgG λ . Po výběru nejlepších proteinů byla cílem jejich produkce, purifikace a základní biofyzikální charakterizace.

Motivací pro *in silico* část této práce bylo získání atomistického pohledu na vazbu v komplexech MLA-10E8 s experimentálně zkoumanými MLA proteiny, vysvětlení mechanismu mimikování vazby MPER-10E8 a získání fyzikálně-chemických údajů o vazebných silách. Součástí práce bylo vyvinutí protokolu, který by dokázal poskytnout teoretické výsledky o dostatečné spolehlivosti i za absence ověřujících experimentálních dat.

4 Materiál a metody

4.1 *In vitro* Experimentální část

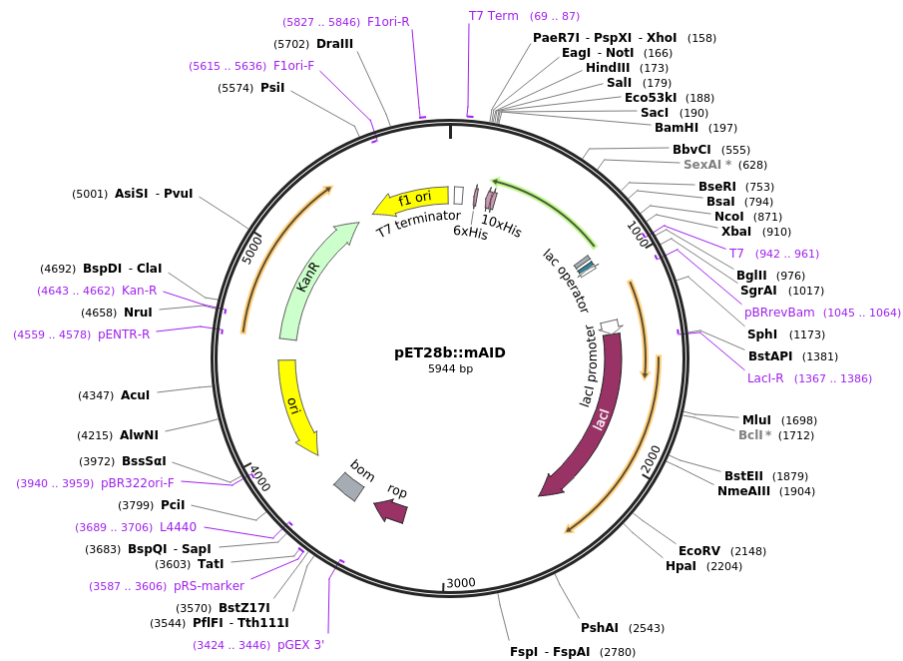
4.1.1 Použité přístroje a materiál

4.1.1.1 Přístroje

Název a popis	Výrobce	Popis
0,22 um filtr + aparatura na vakuovou filtraci		
Allegra X-22R Centrifuge	Beckman Coulter	Centrifugace 15/50 ml zkumavek
Avanti J-26S XPI	Beckman Coulter	Ultracentrifugace
Bioer ThermoCell MB-102	Bioer	Inkubace mikrozkušavek pro transformace a SDS-PAGE
Bio-Rad PowerPac Basic	Bio-Rad	Zdroj pro SDS-PAGE elektroforézu
Boeco BPS 51	Boeco	Laboratorní váha
Chirascan Plus	Applied Photophysics	CD Spektrometr
Epoch 2 Microplate spektrofotometr	BioTek	Spektrometr na čtení ELISA desek
GE AKTA Purifier 10 FPLC	GE	HPLC Chromatograf
Herolab UVT-20M	Herolab	UV iluminátor gelů
IKA RCT basic	IKA	Magnetická míchačka s regulací teploty
Incu-Shaker 10LR	Benchmark	Inkubátor pro větší objemy
Labnet Prism R Microcentrifuge	Labnet	Centrifugace mikrozkušavek
Microfuge 22R centrifuga	Beckman Coulter	Centrifuga na mikrozkušavky s chlazením
Misonix 3000 + sondy	Misonix	Lýze buněk ultrazvukem
Monolith microscale thermophoresis NT.115	Nano Temper	MST s fluorescenčním značením
MR-1 Mini Rocker-Shaker	Biosan	Orbitální třepačka
NanoDrop One	Thermo Scientific	UV-VIS kapkový spektrometr
N-Biotek Shaking Incubator NB-205	N-Biotek	Inkubace menších objemů média s třepáním
Nunc-Immuno MicroWell Polysorp deska	Sigma-Aldrich	96 jamková deska pro ELISA s plochým dnem
Nuve EN120	Nuve	Inkubace Petriho misek
P-Lab PV-1	P-Lab	Vortex pro míchání
Prometheus NT.48	Nano Temper	DSF Spektrometr
Superdex 200 10/300 GL	Sigma-Aldrich	Kolona pro gelovou permeační chromatografii
Tecan HydroFlex	Tecan	Promývání 96 jamkových desek
Vivaspin 4 5000 MWCO	Sartorius	Kolony pro koncentrování proteinů
WPA Biowave II	Biochrom	Spektrometr
Zetasizer Nano ZS90	Malvern Panalytical	DLS Spektrometr
Zetasizer Ultra	Malvern Panalytical	MADLS Spektrometr

4.1.1.2 Protilátky, enzymy, vektory a buňky

Název	Výrobce	Koncentrace / specifikace
Anti-HIV-1 gp41 protilátka 10E8	monoklonální NIH AIDS Reagent Program (#12294)	1 mg/ml
Anti-V5 tag s HRP konjugovaná polyklonální protilátka	AbCam	1 mg/ml
BamHI-HF restriktasa	NEB	20000 jednotek / ml
BL21(λDE3) <i>E.coli</i> kompetentní buňky	Vlastní	
BL21(λDE3)pBirA <i>E.coli</i> kompetentní buňky s birA genem	Vlastní	
DNA Ligasa T4	NEB	
Lidské IgG1 s λ lehkým řetězcem	Sigma-Aldrich	1 mg/ml
NcoI restriktasa	NEB	10000 jednotek / ml
pET-28b vektor	Novagen	Mapa sekvence uvedena na obrázku 9.
XL-1 Blue <i>E.coli</i> kompetentní buňky	Vlastní	V pufru PBS, 1 M MgCl ₂ , 1 M MgSO ₄ ,



Obrázek 9 – Mapa sekvence plasmidu pET28b. Převzato.⁵⁴

4.1.1.3 Pufry, média a ostatní roztoky

Název	Výrobce/specifikace
5× konc. redukující roztok pro SDS-PAGE	10 % w/v SDS; 20 % v/v glycerol; 0,2 M Tris-HCl pH 6,8; 0,05 % w/v bromfenolová modř; 10 mM β-merkapt ethanol
6× konc. nanášecí roztok pro DNA elektroforézu	3 ml glycerol; 3 ml 0,5 M EDTA; 3 mg bromfenolová modř; 3 mg xylenocyanol; 4 ml H ₂ O
10× Cutsmart pufr	NEB
10× T4 DNA Ligase pufr	NEB
Agarové živné médium na misky	35 g agarosy v 1 l LB média.
Akrylamidový zásobní roztok	200 ml, 60 g akrylamid, 1,6 g bis-akrylamid
Anodový pufr	24,22g Trizma-base na 1 l; úprava HCl do pH 8,9
Biotin	5 mM D-Biotin rozpuštěný v 10 mM Bicinovém pufru, pH 8,3
Bradfordův roztok	25 ml EtOH; 50 mg Coomassie Brilliant Blue G-250, 50 ml konc. H ₃ PO ₄ ; do 500 ml H ₂ O; filtrace složeným filtrem
Coomassie R-250 barvivo	50 mg Coomassie Brilliant Blue G-250; 3 ml konc. HAc; do 1 l H ₂ O; filtrace přes papírový filtr
Chloramphenicol zásobní roztok	30 mg/ml v EtOH
Ethidiumbromid	1 % vodný roztok
Kanamycin zásobní roztok	60 mg/ml
Katodový pufr	12,1 g Trizma-base; 17,9 g tricín; 1 g SDS do 1 l vody
LB médium	1 % Trypton; 0,5 % kvasinkový extrakt; 1 % NaCl; 0,1 % D-glukosa; pH 7,0 upravené pomocí NaOH
LB+kan	LB médium + 60 µg · ml ⁻¹ kanamycin
Ni-NTA matrice (Qiagen Ni-NTA agarose)	Qiagen
PBS pufr	137 mM NaCl; 2,7 M KCl; 10 mM Na ₂ HPO ₄ ; 1,8 M KH ₂ PO ₄ ; pH 7,4
PBS-T pufr	PBS pufr s přídáním 0,05 % (w/v) Tween-20; pH 7,4
PBS-T-BSA pufr	PBS-T pufr s přídáním 1 % (w/v) BSA; pH 7,4
Precision Plus Protein™ All Blue Prestained proteinový standard	Bio-Rad
Precision Plus Protein™ Dual Color proteinový standard	Bio-Rad
SDS-PAGE gel	12,5 % spodní gel (1,6 ml H ₂ O; 2,07 ml zás. akrylamidový roztok; 1,25 ml 1,5 M Tris-HCl pH 8,8; 50 µl 10 % SDS; 12,5 µl 25 % APS; 12,5 µl TEMED) Zaostřovací svrchní gel (1,23 ml H ₂ O; 0,27 ml zás. akrylamidový roztok; 0,2 ml 1,5 M Tris-HCl pH 8,8; 60 µl 10 % SDS; 20 µl 25 % APS; 5 µl TEMED)
TAE pufr	4,84 g Tris-base; 1,142 ml konc. HAc; 20 ml 0,5 M EDTA; do 1 l H ₂ O.
TMB-Complete 2 substrát pro HRP reakci	TestLine Clinical Diagnostics s.r.o.
TN Pufr	Tris-base 300 mM; NaCl 50 mM; pH 8,0 upravené HCl
TNI ₂₅₀ eluční pufr	TN pufr s přídavkem 250 mM imidazolu
TNI ₂₀ promývací pufr	TN pufr s přídavkem 20 mM imidazolu
Vazebný pufr ELISA	100 mM Na ₂ CO ₃ /NaHCO ₃ , pH 9,6 upravené pomocí NaOH
Zastavovací roztok ELISA	2 M H ₂ SO ₄

4.1.1.4 Kity

Název	Výrobce
QIAprep Spin Miniprep Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen

4.1.1.5 Internetové servery, aplikace

QtiPlot-0.9.8.9⁵⁵

Webový server ExPASy ProtParam⁵⁶

Webový server UniProt⁵⁷

4.1.1.6 Proteinové sekvence

His-MLA092-V5

MGSSHHHHHSSGKSELAVEILGKGQVRFWMQAEIMWGNKVNIFNEKEIFEGPKYKMHIDPSWGIIEMFMEKLQDEDEGTYTFQLQIVTPLNHST
VVLVGDVFKKLQKEAEFQRQEWIRKQGGSGKPIPPLLGLDST

E. coli Hsp70 / DnaK chaperon

MGKIIGIDLTTNSCAVIMDGTTPRVLENAEGDRTPSIIAYTQDGETLVGQPAKRQAVTNPQNTLFAIKRLIGRRFQDEEVQRDVSIMPFKIIAADNGDA
WVEVKGQKMAPPQISAEVLKMKKTAEDYLGEPTVEAVITVPAYFNDAQRQATKDAGRIAGLEVKRIINEPTAAALAYGLDKGTGNRTIAVYDLGGGTFD
ISIEIDEVDGEKTFEVLATNGDTHLGGEDFDSRLINYLVEEFKQDQIDLRNDPLAMQRLKEAAEKAKIELSSAQTDVNLPLYITADATGPKHMNIKVTRAK
LESLVEDLVNRSIEPLKVALQDAGLSVSDIDDVILVGGQTRMPMVQKVAEFFGKEPRKDVNPDEAVAIGA AVQGGVLTGDVKDVLVLLDVTPLSLGIETM
GGVMTTLIAKNTTIPTKHSQVFSTAEDNQSAVTIHVLQGERKRAADNKSLGQFNLDGINPAPRGMPQJEVTFDIDADGILHVS AKDKNSGKEQKITIKASS
GLNEDEIQKMVRDAEANA EADRKFEELVQTRNQGDHLLHSTRKQVEEAGDKLPADDKTAIESALTALETALKGEDKAAIEAKMQELAQVSQKLEIAQQ
QHAQQQTAGADASANNAKDDDDVDAEFEEVKDKK

4.1.2 Postup práce

4.1.2.1 Transformace bakteriálních buněk

K transformaci kompetentních buněk *E. coli* byl použit vždy standardní postup, kde k 25 μl buněčné suspenze byl přidán 1 μl roztoku plasmidové DNA o přibližné koncentraci $100 \text{ ng} \cdot \mu\text{l}^{-1}$ (v pufru 10 mM TRIS, pH 8,0) nebo 5 μl DNA po ligaci ke 100 μl buněčné suspenze *E. coli* XL1-Blue. Tato směs byla ponechána po dobu 15 minut v drceném ledu. Po této inkubační periodě proběhl teplotní šok vložením zkumavky s transformační směsí do předehřátého termobloku o teplotě $42 \text{ }^{\circ}\text{C}$ po dobu 45 s. Následně byla zkumavka vyjmuta a ponechána 5 min v drceném ledu, poté bylo přidáno alespoň devítinásobné množství média LB, typicky 900 μl na 100 μl buněčné směsi, v případě menších alikvot 500 μl LB na 25 μl buněčné směsi a směs byla inkubována po dobu 1 h za stálého třepání při teplotě $37 \text{ }^{\circ}\text{C}$. Po této inkubaci byla směs rozetřena na agarovou misku s kanamycinem (cca 250 μl buněčné směsi na misku). V případě většího množství transformací variant určených pro produkci v BL21(λ DE3)

či BL21(λ DE3)pBirA byla transformační směs centrifugována při $15\,000 \times g$ po dobu 1 min. Byla odstraněna část supernatantu a ve zbylém objemu (cca 50 μl) byla bakteriální peleta suspendována a nanesa na agarovou misku s kanamycinem ($60 \text{ } \mu\text{g} \cdot \text{ml}^{-1}$) a chloramfenikolem ($30 \text{ } \mu\text{g} \cdot \text{ml}^{-1}$). Agarové misky byly ponechány inkubaci přes noc při $37 \text{ }^{\circ}\text{C}$.

Plasmidová DNA kódující His-MLA-Avi konstrukty MLA102 a 139-190 byla transformována do BL21(λ DE3)pBirA bakteriálních buněk.

4.1.2.2 Produkce proteinů pro plošnou selekci

Bakteriální kolonie z agarové misky byly přeneseny do 0,5 ml LB média ve 2 ml zkumavkách a suspendovány mírným vortexováním. Po dobu 2 h byly inkubovány za stálého třepání při teplotě $37 \text{ }^{\circ}\text{C}$. 100 μl suspenze bylo přeneseno do 5 ml LB média s kanamycinem v 50 ml centrifugačních zkumavkách, následovala 2h inkubace. Pro indukci exprese biotinylovaného proteinu byl připraven roztok obsahující 5 mM biotin a 100 mM IPTG, tento roztok byl přidán do kultur po 2 hodinách inkubace v objemu 55 μl

do 5 ml kultury (typicky 50 μ l 5mM biotinu a 5 μ l 1M IPTG na 5 ml kulturu). Pro zvýšené provzdušnění byla víčka zkumavek částečně povolena. Bakteriální kultury byly třepány při 22 °C přes noc. Zkumavky s kulturami byly centrifugovány při 4500 \times g, 4 °C po dobu 15 min.

Byl odstraněn supernatant a pelety byly suspendovány v 5 ml TN pufru. Buňky byly lyzovány sonikací (Qsonica, USA) po celkovou dobu 1 min, kde se cyklicky střídalo 20 s pulsu a 10 s chlazení při výkonu 5W za současného chlazení zkumavky drceným ledem a vodou. Suspenze byla centrifugována po dobu 20 min při 40 000 \times g v Avanti centrifuze.

4.1.2.3 SDS-PAGE elektroforéza

Obecný postup SDS-PAGE elektroforézy byl následující. K 1-15 μ l vzorku byl vždy přidán 5 \times ředěný redukující roztok pro SDS-PAGE. Vzorky byly po dobu 10 min zahřívány při 95 °C v termobloku, následně byly injekcí přeneseny do jamek SDS-PAGE gelu. Do jedné z jamek byl přidán 1 μ l standardu Precision Plus Protein (buď All Blue, nebo Dual Color verze). Vnitřní strana aparatury byla naplněna po okraj katodovým pufrům, vnější strana pak po rysku anodovým pufrům. Elektroforéza byla prováděna při konstantním napětí 150 V.

Po provedení elektroforézy byly gely vyňaty z aparatury a v destilované vodě zahřívány mikrovlnou troubou do dosažení varu. Voda byla vyměněna a postup byl zopakován, celkově proběhly vždy 3 opakování. Gely byly ponořeny do Coomassie R-250 barviva, znovu povařeny a ponechány třepání na orbitální třepačce do uspokojivého vybarvení. Barvivo bylo nahrazeno vodou a gel byl opět 3 \times povařen. Po posledním povaření byl ponechán ve vodě a třepán na orbitální třepačce do uspokojivého odbarvení. Takto získaný obarvený gel byl skenován či vyfocen.

Ze všech roztoků proteinů určených pro plošnou selekci bylo odebráno 5 μ l, přidáno bylo po 1 μ l redukujícího barvicího roztoku.

4.1.2.4 ELISA, plošná selekce z proteinové knihovny

Popsán je obecný postup metody ELISA, společný všem analýzám. Imunoadsorpční destička NUNC Polysorp (96 jamek) byla pokryta protilátkami, 10E8 IgG a kontrolní izotopovou protilátkou IgG λ , naředěnými ve vazebném pufru do výsledné koncentrace $5 \mu\text{g} \cdot \text{ml}^{-1}$ tak, aby na každou používanou jamku desky připadalo 80 μl roztoku. Dle schématu experimentu bylo vždy nanášeno po 75 μl těchto roztoků do korespondujících jamek. ELISA deska byla ve tmě a při pokojové teplotě ponechána inkubaci po dobu 2 h. Deska byla promyta roztokem PBS-T pomocí promývačky Tecan HydroFlex. Následovalo blokování neobsazených vazebných míst jamek pomocí 300 μl PBS-T-BSA a 2 h inkubace při pokojové teplotě za nepřístupu světla. Deska byla promyta. Poté byly nanášeny analyzované proteiny v roztocích o objemech 75 μl . Případná kontrola vazby na BSA byla v tomto kroku provedena vynecháním přídavku proteinu do konkrétní jamky. Po 1 h inkubace a promytí byla nanášena detekční protilátka či streptavidin-HRP, vždy 10 000 x ředěná v PBS-T-BSA a přenesena do všech jamek po 75 μl . Následovala 1 h inkubace a promytí.

Detekce probíhala spektrometricky. Bylo přidáno 100 μl TMB-Complete 2 substrátu, poté následovala 5-15 min inkubace (v závislosti na síle signálu). Barvicí reakce byla zastavena 2 M roztokem H_2SO_4 a byly odečteny signály jednotlivých jamek (absorbance při $\lambda = 450 \text{ nm}$) pomocí Epoch 2 spektrometru. Statistická analýza a vizualizace proběhla pomocí softwaru QtiPlot.⁵⁵

Pro účely plošné selekce byly proteiny pro třetí krok ELISA analýzy připraveny tak, že bylo vždy 15 μl supernatantu získaného centrifugací ředěno přenesením do 500 μl PBS-T-BSA, každý vzorek lyzátu byl pipetován v tripletech po 75 μl na jamku včetně kontrolních jamek s izotopovou kontrolou IgG λ a negativní kontrolou BSA.

Na základě získaných poměrů mezi vazbou na 10E8 a kontrolní protilátku IgG λ byly vybrány nejlepší klony, které byly identifikovány sekvenací jejich plasmidové DNA (servisní středisko BIOCEV, PřF UK) a takto vyselektované sekvenční varianty byly dále charakterizovány.

4.1.2.5 Záměna C-terminálního Avi-tagu za V5-tag

Plasmidy kódující varianty Myomedinu MLA014, MLA015, MLA016, MLA024, MLA025, MLA045, MLA072, MLA079, MLA092, MLA093, MLA132, MLA158, MLA159 v konstruktech His-MLA-Avi byly vloženy (viz **kap. 4.1.2.1**) do buněk *E.coli* XL1-Blue. Získané bakteriální kultury byly odebrány a přeneseny do 5 ml LB+kan a pěstovány za stálého třepání přes noc při 37 °C. Získaná bakteriální kultura byla použita pro extrakci plasmidové DNA pomocí QIAprep Spin Miniprep Kit Kitu podle příloženého protokolu. Eluce byla provedena přidáním 50 µl elučního pufru (součástí kitu).⁵⁸ Pomocí spektrometrického měření absorbance při $\lambda = 280$ nm a $\lambda = 260$ nm na Nanodrop spektrometru byly naměřeny koncentrace jednotlivých izolovaných plasmidů.

Pro štěpení DNA byly připraveny reakční směsi obsahující 27 µl plasmidové DNA, 3,1 µl 10× CutSmart pufru, 0,5 µl roztoku restriktasy NcoI a 0,5 µl roztoku restriktasy BamHI-HF. Směsi byly inkubovány po dobu 2 h při teplotě 37 °C v inkubátoru. Štěpené směsi (31 µl) byly smíchány s 5 µl nanášecího roztoku pro DNA a byly pipetovány do jamek 1 % agarosového gelu s přidavkem ethidiumbromidu (2 µl 1% vodného roztoku na 50 ml gelu). Elektroforéza probíhala po dobu 1 h při konstantním napětí 80 V.

Na základě pohyblivosti v gelu prosvíceném UV zářením byly vybrány proužky odpovídající velikosti insertové DNA. Ty byly vyříznuty skalpelem a DNA z gelu byla extrahována pomocí QIAquick Gel Extraction kitu dle příloženého návodu.⁵⁹ Opět byly spektrometricky naměřeny koncentrace DNA dříve zmíněnou metodou.

Do ligační reakce o celkovém objemu 20 µl bylo pipetováno 2 µl 10× T4 DNA Ligase pufru, 2 µl štěpeného plasmidu pET-28b (koncentrace obvykle kolem 30-50 ng · µl⁻¹) s V5 tagem na 3'-konci expresní kazety, 1 µl roztoku DNA Ligasy T4, a izolovaná insertová DNA do celkového objemu 20 µl (koncentrace obvykle kolem 20-30 ng · µl⁻¹). Směs byla ponechána po dobu 1 h při pokojové teplotě. Paralelně byla provedena i kontrolní ligace, kde insertovou DNA nahradil eluční pufr.

K transformaci buněk *E.coli* XL-Blue bylo použito 5 μ l ligačních směsí dle dříve uvedeného protokolu. Pro každou variantu byly odebrány z Petriho misky dvě kolonie a pěstovány v 5 ml LB+kan média při 37 °C přes noc za neustálého třepání při 220 RPM. Ze získaných saturovaných bakteriálních kultur byla izolována plasmidová DNA, která byla ověřena sekvenováním (Servisní středisko BIOCEV, PřF UK) a dále sloužila jako zdroj kódující konstrukty His-MLA-V5 především pro transformaci expresních kmenů *E.coli* pro produkci proteinu.

4.1.2.6 Produkce rekombinantních proteinů pro analýzy vazebné afinity

Obecný postup produkce proteinů pro jejich charakterizaci byl následující. Pomocí plasmidů kódujících konstrukt His-MLA-V5 byla provedena transformace BL21 (DE3) buněk, které byly selektovány na agarových miskách s kanamycinem. Po inkubaci přes noc při 37 °C bylo vždy odebráno několik kolonií, které byly přeneseny do 0,5 ml LB média a suspendovány pipetováním. Suspenze byly inkubovány po dobu 2 h ve 37 °C za stálého třepání. Celý objem bakteriální kultury byl přenesen do 200 ml LB+kan média v Erlenmeyerově baňce. Ta byla uzavřena prodyšným víčkem a třepána při 37 °C v inkubátoru. Během této doby byla sledována každých 30 min optická densita při $\lambda = 600$ nm (OD_{600}) oproti standardu tvořenému čistým LB+kan médiem. Měření proběhlo na přístroji WPA Biowave II.

Po dosažení $OD_{600} \sim 0,6$ bylo do kultur přidáno 100 μ l zásobního roztoku 1 M IPTG pro indukci exprese. Ta probíhala přes noc při 22 °C. Kultury byly centrifugovány při 4500 \times g po dobu 15 min při 4 °C. Supernatant byl odstraněn a k peletě bylo přidáno 15 ml TN pufru. Lýze buněk proběhla sonikací za současného chlazení zkumavky v drceném ledu a vodě při amplitudě 2,5 odpovídající výkonu 15W, po dobu 5 min (čistý čas sonikace, celkem 15 min včetně chlazení) s cykly 20 s sonikace a 40 s chlazení. Sonikovaný lyzát buněk byl centrifugován pomocí vysokorychlostní centrifugy Avanti při 40 000 \times g po dobu 20 min a chlazení při 4 °C a supernatant byl použit na purifikaci proteinu afinitní chromatografií na Ni-NTA sefarosové matrici.

Takto byly produkovány varianty MLA014, MLA015, MLA016, MLA024, MLA025, MLA045, MLA072, MLA079, MLA092, MLA093, MLA132, MLA158, MLA159 a MyoWT.

4.1.2.7 Purifikace na Ni-NTA agarose

Do 15 ml plastových kolon určených pro gravitační chromatografii bylo naneseno po 2 ml suspenze matrice Ni-NTA Agarosa (QIAGEN). Po průtoku kapalně fáze fritou kolony byla matrice ekvilibrována několikanásobným promytím 5 ml TN pufru. Celý objem centrifugovaných lyzátů (supernatantu) získaných sonikací byl přenesen na kolonu a průtok byl sbírán zpět do zkumavky. Toto bylo provedeno celkem 3x.

Byla připravena 96-jamková mikrotitrační destička, do jejíchž jamek bylo pipetováno po 90 μ l Bradfordova činidla. Matrice byla následně promývána TN_{I20} puftrem a průtok byl sbírán vždy po 10 μ l, které byly přenášeny do jamek 96-jamkové desky a bylo pozorováno zeslabující modré zbarvení indikující absenci proteinu v protékajícím objemu. Po ztrátě zbarvení Bradfordova činidla průtokem bylo promývání ukončeno.

Samotná eluce probíhala vytěsňováním pomocí TN_{I250} pufru v 0,5 ml objemech do připravených mikrozkušavek. Obdobně jako v kroku promývání bylo pozorováno zbarvení Bradfordova činidla a po ztrátě pozorovaného zbarvení způsobeného 10 μ l odebraného vytékajícího eluátu byla eluce vždy ukončena. Jednotlivé frakce včetně průtoků z promývacích kroků byly analyzovány pomocí SDS-PAGE.

Všechny v předchozím kroku pěstované varianty byly purifikovány pomocí afinitní chromatografie na Ni-NTA agarose.

4.1.2.8 Analýza vazebné afinity metodou ELISA a příprava pro *in vivo* experimenty

Pro každý roztok purifikovaných MLA proteinů včetně MyoWT byla spektrometricky naměřena absorbance při $\lambda = 280$ nm. K výpočtu přibližné koncentrace proteinu v roztoku bylo využito teoretických extinkčních koeficientů získaných pomocí webového serveru ExPASy ProtParam.⁵⁶ Molekulová hmotnost všech konstruktů byla uvažována na 16 kDa. Imobilizace protilátek 10E8 a IgG λ na imunoabsorpční destičky proběhly za identických podmínek popsanych **kapitolou 4.1.2.4**.

Pro každou variantu byla následně připravena osmi bodová ředící řada v PBS-T-BSA pufru, jednotlivá ředění byla na destičku nanášena v dubletech, jeden dublet pro imobilizovaný 10E8 a druhý dublet pro IgG λ .

V prvním bodě ředící řady odpovídajícím nejvyšší koncentraci proteinu byly všechny varianty ředěny tak, aby maximální koncentrace proteinu odpovídala jednotkám mM. Pak následovalo vždy ředění 5 \times v PBS-T-BSA pufru. Vzniklé roztoky variant získané ředící řadou byly naneseny na desku.

Takto připravená deska byla inkubována, byly přidány protilátky Anti-V5 a bylo provedeno barvení, detekce a statistická analýza výsledků. Hodnoty K_d byly zjištěny proložení průběhu absorbance/koncentrace logistickou funkcí o 4 parametrech ($y = \max + \frac{\min - \max}{1 + \left[\frac{x}{x_0}\right]^p}$), kde *max* a *min* označují maximum a minimum funkce; x_0 označuje polohu inflexního bodu sigmoidy; p označuje Hillův parametr/sklon). Proložení proběhlo v aplikaci QtiPlot pomocí maxima 1000 iterací, „Scaled Levenberg-Marquardt“ algoritmem s tolerancí 0,0001).

Detailní postup experimentů ELISA popisuje **kapitola 4.1.2.4**.

Pro *in vivo* testování byly pěstovány varianty MLA016, MLA024, MLA025, MLA045, MLA072, MLA079, MLA092, MLA093, MLA132, MLA158 a MLA159. Pěstování proběhlo v 500 ml média dle dříve popsaného postupu.

4.1.2.9 Purifikace SEC

Obdobnou metodou, jaká je zmíněna v kapitole 4.1.2.6 byly pěstovány varianty MLA024, MLA025, MLA092 a MLA158 v podobě konstruktů His-MLA-V5. Obdobně byl připraven i protein His-MyoWT-V5. Varianty MLA024 a MLA025 v objemu 500 ml LB+kan, MyoWT v objemu 200 ml, MLA092 v 2400 ml. Získané bakteriální pelety byly vždy suspendovány v 30 ml TN_I₂₀ pufru a sonikovány. Byl odebrán 20 µl vzorek tohoto bakteriálního lyzátu. Takto vzniklých 15 ml směsi bylo centrifugováno při 40 000×g při 4 °C po dobu 15 min.

Purifikace probíhala dvoukrokově, afinitní chromatografií následovanou gelovou chromatografií. Prvním krokem byla purifikace pomocí Ni-NTA Agarosy (4.1.2.7), prováděná se změnou oproti předchozímu postupu. Kolona byla vždy ekvilibrována pufrům TN_I₂₀, identickým pufrům, ve kterém byly i roztoky proteinů určené pro purifikaci. Všechny frakce včetně průtoku po nanášení na kolonu byly uschovány v objemech 20 µl. Byly provedeny SDS-PAGE elektroforézy.

Všechny eluované frakce proteinů byly spojeny a byl přidán roztok EDTA do výsledné koncentrace 4 mM. Spojený roztok byl centrifugován při zrychlení 18 000×g po dobu 10 min při 4 °C. Odebraný supernatant byl koncentrován pomocí ultrafiltrace s centrifugací na Vivaspin 4 5000 MWCO kolonách při 9800×g, 4 °C po dobu 10 min. Po každém stočení byl roztok promíchán opatrným otáčením kolony. Tento proces byl opakován, dokud celkový objem roztoku v horní části kolony nebyl přibližně 500 µl nebo méně. Jak koncentrát, tak průtok ve spodní části kolony byly analyzovány metodou SDS-PAGE (5 µl koncentrátů, 20 µl průtoku).

Získaný koncentrovaný protein byl přenesen do mikrozkušavky, ve které byl centrifugován 10 min při 18 000×g po dobu 10 min při 4 °C. Objem zkumavky byl zjištěn přibližně vážením v porovnání s prázdnou mikrozkušavkou stejného typu. Byla měřena absorbance při $\lambda = 280$ nm (oproti průtoku koncentrátoru) na základě kterých ve spojení s teoretickými extinkčními koeficienty byly vypočteny hmotnostní koncentrace koncentrovaných roztoků proteinů.

Druhým krokem purifikace byla gelová permeační chromatografie s využitím GE AKTA Purifier 10 systému a Superdex 200 10/300 kolony.

Bylo připraveno 500 ml 0,22 μm filtrem přefiltrované, odplyněné destilované vody a PBS pufru. Filtrace i odplynění probíhalo pomocí aparatury pro vakuovou filtraci za stálého míchání magnetickou míchačkou s míchadlem po dobu 1,5 hodiny za mírného zahřívání ($\sim 40\text{ }^\circ\text{C}$). Odplyněnou filtrovanou vodou byla promyta Superdex kolona (36 ml, průtok $1,0\text{ ml}\cdot\text{min}^{-1}$) včetně pumpy. Následně byla stejně promyta kolona a pumpa odplyněným roztokem PBS. 0,5 ml nástřiková smyčka byla manuálně promyta pomocí 10 ml odplyněného PBS z injekční stříkačky (pozice LOAD).

Hamiltonovou stříkačkou byly nanесeny vzorky do nástřikové smyčky během stálého průtoku PBS rychlostí 0,5 ml/min. Vzorek ze smyčky byl přesměrován do toku na kolonu. Pomocí UV detektoru byla snímána absorbance při $\lambda = 280\text{ nm}$ celou dobu průtoku. Frakce byly sbírány automatickým kolektorem každého 0,5 ml (po 6 ml mrtvého objemu kolony) nebo při prudkém nárůstu či poklesu absorbance. Dle chromatogramů byly sesbírány frakce obsahující protein. Kolona i aparatura byla promyta destilovanou vodou, roztokem 20 % ethanolu.

Získaná data z chromatografie byla vynesena pomocí aplikace QtiPlot a Matplotlib knihovny jazyka Python.

Na sesbíraných frakcích byla provedena SDS-PAGE elektroforéza. Byly měřeny absorbance vzorků při $\lambda = 280\text{ nm}$ a na jejich základě (s předpokladem stejného teoretického absorpčního koeficientu u všech frakcí) byly vypočteny koncentrace proteinů. Frakce odpovídající jednotlivým chromatografickým píkům byly zakonzentrovány ultrafiltrací (viz dříve) do výsledného objemu přibližně 500 μl . Tři charakteristické píky byly označeny dle elučního času sestupně jako SEES, SEMS a SELS (size-exclusion-early/middle/late series). Toto značení bude dále užíváno.

4.1.2.10 ELISA jednotlivých SEC separovaných frakcí

Pro jednotlivé frakce (varianty MLA024 a MLA025) či zakoncentrované roztoky odpovídající frakcím chromatografických píků (MLA092, MyoWT) bylo provedeno stanovení vazebných afinit pomocí metody ELISA. Zároveň bylo provedeno porovnání se vzorky MLA024 a MLA025 purifikovanými pouze pomocí Ni-NTA.

Konkrétní použitá ředění pro jednotlivé roztoky proteinů jsou uvedena v **tabulce 3**, která zároveň zobrazuje, které SEC získané frakce byly zakoncentrovány a využity v dalších experimentech.

Tabulka 3 – Získané frakce MLA proteinů po SEC purifikaci, jejich zakoncentrování (pravý sloupec) a ředění pro ELISA experimenty.

	$V_{\text{protein}}; V_{\text{pufr}} (\mu\text{l})$	$V_{\text{pufr}}; V_{\text{přenos}} (\mu\text{l})$	SEC frakce
MLA024 Ni-NTA	5,94 ; 422	320; 106,6	X
MLA025 Ni-NTA	15,0 ; 412,3	320; 106,6	X
MLA024 SELS	133,2 ; 267	320; 80	5
MLA025 SELS	285,6 ; 144,2	320; 80	5
MLA024 SEMS	400 ; 0	320; 80	9
MLA025 SEMS	200; 200	320; 80	9
MLA024 SEES	138,8; 261	320; 80	24
MLA025 SEES	106,6; 293,3	320; 80	23
MLA092 SEES*	**	**	3+4
MLA092 SEMS	35,6; 284	320; 106,7	9 až 12
MLA092 SELS	160; 160	320; 160	18 až 25
MLA158 SEES	**	**	4 až 6
MLA158 SEMS	**	**	8 až 10
MLA158 SELS	**	**	18 až 22
MyoWT SEES*	**	**	4+5
MyoWT SEMS	**	**	7 až 9
MyoWT SELS	**	**	18 až 24

* po zakoncentrování nebyl získán protein ** nebyl proveden experiment ELISA

4.1.2.11 Biofyzikální metody – MS, CD, DLS, MST, DSF

Zakoncentrované frakce proteinů z SEC popsané v kapitole 4.1.2.10 byly charakterizovány biofyzikálními metodami.

Frakce MLA024 SELS byla společně s vzorkem purifikovaným pomocí Ni-NTA agarosy MyoWT podrobena měření cirkulárního dichroismu oproti čistému PBS. (provedeno Dr. Tatsianou Chernavets v CMS Biocev) na přístroji Chirascan Plus. Měření proběhlo v roztoku s pufrům PBS s proteiny o přibližné koncentraci $0,2 \text{ mg} \cdot \text{ml}^{-1}$.

Pro dva z SDS-PAGE gelu SEC purifikovaných frakcí His-MLA092-V5 skalpelem vyříznuté pruhy proteinů byla provedena LC-MS (MALDI-TOF) analýza s cílem identifikace proteinů. K analýze sekvencí byly využity dvě známé sekvence, odpovídající His-MLA092-V5 a bakteriálnímu Hsp70 proteinu (získané pomocí serveru UniProt)⁵⁷. Samotné měření nebylo předmětem práce autora.

Získané frakce byly měřeny metodou dynamického rozptylu světla (DLS) na homogenitu. Měření proběhlo v křemenných květech v objemu $40 \mu\text{l}$ na přístrojích Zetasizer Nano ZS90 a Zetasizer Ultra v „backscatter“ módu po 120 s temperování aparatury na $25 \text{ }^\circ\text{C}$.

Ve skleněných kapilárách proběhlo měření termostability přístrojem Prometheus NT.48 metodou diferenční skenovací fluorimetrie pomocí tryptofanů. Byl využit teplotní gradient začínající na $20 \text{ }^\circ\text{C}$, vzrůstající o $1,5 \text{ }^\circ\text{C}$ každou minutu do maximální teploty $95 \text{ }^\circ\text{C}$. Intenzita excitačního záření byla zvolena tak, aby byla fluorescenční emise v mezích měření. Měření probíhalo ve skleněných kapilárách s objemy zhruba $4 \mu\text{l}$.

4.2 *In silico* část

4.2.1 Proteinové sekvence

Všechny sekvence popisované v této práci sestávají z N-terminálního 6×-His-Tagu, samotné myomedinové sekvence a C-terminálního V5-tagu s krátkým linkerem GS.

6×-His-Tag (13 AK): MGSSH HHHHH SSG
Myomesin (111 AK): KSELA VEILE KGQVR FWMQA **XXXX**G NAKVN YIFNE KEIFE
GPKYK MHID**X XX**GII EMFME KLQDE DEGTY TFQLQ **XXXXX**
NHSTV VLVGD VFKKL QKEAE FQRQE WIRKQ G
Linker-V5-Tag (16 AK): GS-GKP IPNPL LGLDS T

Konkrétní aminokyseliny na červeně vyznačených pozicích jsou pro jednotlivé varianty uvedeny v **tabulce 4**.

Tabulka 4 – Aminokyselinové sekvence vazebných proteinů řady MLA založených na proteinovém scaffoldu Myomedinu. Uvedené číslování pozic odpovídá celému konstruktu včetně tagů 6× His a V5.

Varianta	p.34-37 (klastr 1)	p.63-65 (klastr 2)	p.89-93 (klastr 3)
MLA024	LSVF	RNT	FMLMM
MLA025	ATPS	GHE	VILIL
MLA092	EIMW	PSW	IVTPL
MLA158	WMWW	ITL	LYYAW
MyoWT	EKLS	RNT	DGKAT

4.2.2 Použitý software, skripty a balíky

Vizualizace, homologní modelování, analýza modelů

MODELLER 9.24⁶⁰

MolProbity (lokální verze – balíček skriptů)⁶¹

PROPKA3 (lokální verze)⁶²

PyMol 2.0⁶³

UCSF Chimera 1.14⁶⁴

VMD 1.9.3⁶⁵

Molekulová dynamika a silová pole

AmberTools20⁶⁶

GROMACS 2020.3 a 2019.6^{67,68}

CHARMM36m (Červen 2020, GROMACS)⁶⁹

ff99SB⁷⁰ (AMBER)

SIRAH x2.2_20-07 (GROMACS)⁷¹

Analýza vazebných interakcí

g_mmpbsa upravená verze⁷²

GMX_mmpbsa 1.4.2 (balíček skriptů)⁷³

gRINN⁷⁴

g_wham (součást GROMACS)⁷⁵

MMPBSA.py (součást AmberTools20)⁷⁶

Webové servery

ClusPro2.0^{77–81}

GalaxyHomomer¹⁸²

GalaxyWEB⁸³

ModRefiner⁸⁴

PDB2PQR⁸⁵

Visualizace a další

Python, Miniconda3, Matplotlib, Seaborn, Numpy, vlastní knihovna

HYDROPRO⁸⁶

Proteinové struktury

2KHO (PDB)⁸⁷

4G6F (PDB)⁸⁸

3RBS (PDB)⁸⁹

Chaperone protein DnaK (*E. coli*, AlphaFold)^{90,91}

4.2.3 Modelování, selekce a simulace

4.2.3.1 Homologní modelování, vylepšení a validace

Strukturní modely vazebných proteinů MLA s N-terminálním His-tagem a C-terminálním V5-tagem byly získány homologním modelováním s využitím softwaru Modeller skrz rozhraní v UCSF Chimera. Konkrétní sekvence proteinů (viz **tabulka 4**) byly použity pro vyhledání templátů z databáze PDB pomocí algoritmu BLAST.

K vybudování modelů byl použit jeden templát, dostupný v PDB databázi pod kódem 3RBS.⁸⁹ Bylo vytvořeno po pěti modelech na protein, z nichž byl vybrán vždy nejlepší model dle nejvíce negativního zDOPE skóre a dle vizuální inspekce.

V UCSF Chimera byly vygenerovány Ramachandranovy výnosy pro každý z vybraných modelů, které byly použity pro další kontrolu kvality modelů pomocí kontroly dovolených kombinací dihedrálních úhlů ϕ a ψ proteinů. Teoretické hodnoty pKa aminokyselin vybraných modelů a hodnoty pH optim stability byly vypočteny pomocí skriptu ProPka. Pomocí aplikace PyMol byly takto vytvořené homologní modely vizualizovány.

Takto vzniklé homologní modely (dále „prvotní modely“) byly dále podrobeny validaci pomocí MD simulací v balíku GROMACS 2020.3. Časový krok všech MD simulací činil $\delta t = 2 \text{ fs}$ s „constraints“ na vazbách zahrnujících vodík pomocí algoritmu SETTLE.

Struktury byly vloženy do dodekaedrálních boxů se vzdáleností ke krajům boxu minimálně 1.2 nm v systému s PBC. Solvatovány byly pomocí SPC216 modelu vody pod silovým polem CHARMM36m a modelem vody TIP3P. Proběhla neutralizace ionty Na^+ a Cl^- , dále bylo přidáno těchto iontů do výsledné koncentrace 0.150 M. Byla

provedeny dvě kola minimalizace struktur, první pomocí „steepest descent“ algoritmu do bodu, kdy byla splněna podmínka $E < 1000 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-1}$ pro celkovou energii E systému. Následně byla provedena druhá minimalizace algoritmem „conjugate gradient“ do konvergence průběhu energie. Následovala 100 ps ekvilibrace v NVT souboru s couplingem dvou skupin (protein + ostatní částice) s V-rescale (modifikovaný Berendsenův) termostatem s $\tau_t = 0.1$ a cílovou teplotou 300 K. Během této ekvilibrace byly atomy proteinové páteře „restrainovány“ silou $1000 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-1}$. Ekvilibraci předcházelo vygenerování rychlostí částic pomocí Maxwell-Boltzmannova rozdělení. Všechny další MD simulace obsahovaly identický teplotní coupling.

Následovala 100 ps ekvilibrace v NPT souboru s tlakovým couplingem s Parrinello-Rahmanovým barostatem s parametrem $\tau_p = 2.0$ ps a izotermální kompresibilitou vody $4,5 \cdot 10^{-5} \text{ bar}^{-1}$. Referenční tlak byl nastaven na 1 bar, komprese byla isotropaní.

Výsledné průběhy energií systémů, teplot, tlaků a hustot byly použity jako kontrola kvality přípravy systému na produkční MD simulaci.

Produkční simulace byly délky 20 ns. Po těchto simulacích byly zkontrolovány energetické termy (tlak, potenciál, kinetická energie, objem boxu, hustota roztoku, geometrické parametry boxu) se statistickým vyhodnocením ze všech snímků simulace.

Dále byly kontrolovány geometrické parametry jako jsou RMSD, RMSF, gyační poloměr a minimální vzdálenost k periodickým obrazům proteinu. Na základě těchto kontrol byly všechny další kroky práce prováděny pouze na proteinech, s odstraněním terminálních tagů. Z trajektorií byly vybrány konkrétní reprezentativní centroidní konformace, které sloužily jako vstupní struktury do dalších kroků (výběr viz dále).

Kvalita modelů vzniklých z výběru reprezentativních centroidálních struktur trajektorií byla provedena geometrickými kritérii pomocí balíku MolProbity. Sledovanými parametry byly Van der Waalsovské překryvy atomů, rotamery, Ramachandranovy parametry dihedrálních úhlů, atypické vazebné úhly a vzdálenosti. Stejným balíkem byl v případě výhodnosti vzhledem k prováděným testům, tedy pokud takový zásah vylepšil geometrii, v případě MLA024 otočen Asn43 o 180° v ose řetězce aminokyselinového zbytku. Stejná operace byla provedena pro Gln109 MLA025 a pro Gln86 MLA092.

Z několika analyzovaných centroidních struktur (2-5 struktur) byla vybrána vždy ta nejlepší. Stejná kritéria byla použita i pro hodnocení modelů získaných z následujících kroků.

Druhým krokem vylepšování homologních modelů bylo další kolo MD simulací se vstupními strukturami popsanými výše. Simulace byly prováděny identickým protokolem, tentokrát však s “restraints” $1000 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-1}$ aplikovanými na hlavní řetězec proteinu. Po provedení MD simulací byly vybrány průměrné struktury (nikoliv centroidní), které byly minimalizovány pomocí dříve popsaných minimalizačních postupů a opět analyzovány pomocí MolProbity.

V tomto kroku bylo ukončeno vylepšování homologních modelů MLA024 a MLA158 díky dostatečné kvalitě modelů. Poslední krok vylepšování proběhl pouze na variantách MLA25, MLA92 a MyoWT, které měly atypické vazebné délky nebo úhly lokalizované v pozicích aminokyselinových záměn.

Posledním krokem vylepšování modelů bylo „lokální vylepšování“. Byly provedeny MD simulace se zmraženými atomy všech aminokyselin kromě chybných cílených zaměněných aminokyselin a jejich sousedních aminokyselin, mobilní tedy byla vždy specifická aminokyselina a dva její sousedi. Toho bylo docíleno pomocí indexových skupin GROMACSu a rozdělením struktur na dvě části, zmraženou a mobilní. Vstupy jsou popsány **přílohou B 2**. Vodíky byly z mražení vynechány, jsou řešeny standardně algoritmem SHAKE. Geometrie zmražení byla identická ze všech směrů X Y Z.

Z takto provedených nerovnovážných MD simulací byly vybrány opět reprezentativní centroidní struktury, které byly analyzovány obdobně jako v předchozích popsaných krocích. Výsledné vylepšené homologní modely struktur Myomedinových variant MLA024, MLA025, MLA092, MLA158 a MyoWT byly použity v dalších krocích práce.

Pro kontrolu funkčnosti popsaného protokolu byly výstupní modely MODELLERu (prvotní modely) vylepšeny pomocí standardních nástrojů – webových serverů GalaxyWEB a ModRefiner. Na základě evidentního vylepšení popsaným protokolem MD simulací v porovnání s neuspokojivým výsledkem standardních nástrojů byl pro další práce zvoleny již popsané modely.

4.2.3.2 Dokování a prvotní výběr modelů

Byly připraveny vstupní struktury pro dokování. Jako ligandy byly využity homologní modely proteinů MLAXxx, získané v předchozích krocích. Jako receptor byla vybrána vyřešená struktura protilátkového Fab fragmentu 10E8 v komplexu s nativním ligandem MPER, dostupná v databázi PDB pod kódem 4G6F.⁹² Z modelu byly vyjmuty přítomné molekuly vody a MPER ligand, který byl uložen zvlášť. Fab fragment byl následně zkrácen tak, že byly odstraněny konstantní regiony vyjmutím aminokyselin od nestrukturovaných řetězců mezi variabilní a konstantní doménou fragmentu Fab. V těžkém řetězci byly takto odstraněny aminokyseliny Ser113-Lys214, v lehkém pak Gln108-Glu211. Takto vzniklá zkrácená struktura (dále 10E8fr.) byla ověřena pomocí MolProbity. Ve struktuře MPER peptidu byly pomocí MODELLERu opraveny postranní řetězce argininů.

Vstupními strukturami pro dokování tedy byly:

- a) Receptor – upravený Fab fragment protilátky 10E8 s odstraněnými molekulami vody (PDB: 4G6F), dále označován 10E8fr.
- b) Ligandy – homologní modely MLA024, MLA025, MLA092, MLA158, MyoWT, opravená struktura MPER (nativní ligand)

Dokování proběhlo dvěma způsoby:

- i) Webový server ClusPro2 – Ab initio dokování. Skórovací mód „Antibody mode“. Vybráno bylo 10 nejlepších struktur.
- ii) Webový server HDock – Ab initio dokování kombinované s dokováním dle homologních templátů. Vybráno bylo 10 nejlepších struktur.

Protože vstupním receptorem byl zkrácený 10E8fr., bylo u každého modelu zkontrolováno, zda se ligand neváže na nativně neexistující povrch (vzniklý odebráním části proteinu ze struktury).

4.2.3.3 Hrubý výběr nejlepších modelů a analýza vazebné afinity

Deset nejlepších modelů každého proteinu získaných z ClusPro2 dokování bylo podrobena 20 ns simulaci podle dříve uvedeného protokolu. Získané trajektorie byly analyzovány z geometrického a energetického hlediska, jak bylo již popsáno dříve.

Výsledné průběhy geometrických parametrů byly vizualizovány pomocí vlastní Python knihovny, sledovanými kritérii bylo dosažení konvergence základních parametrů jako je RMSD a Rg a absence kontaktů s periodickými obrazy. Na základě „průměrného času dosažení konvergence“ bylo pro další analýzu vybráno posledních 5 ns každé trajektorie (snímky každých 100 ps, celkově 50 snímků).

Hrubý výběr modelů na základě vazebné afinity byl proveden metodou MM/PBSA. Samotné analýzy byly provedeny pomocí dvou aplikací:

- a) Pomocí skriptu MMPBSA.py z balíčku AmberTool20 ve spojení s balíkem gmx_MMPBSA.
- b) Balíčkem g_mmpbsa s pomocí externího balíku APBS, upraveného pro funkčnost v kombinaci s GROMACS 2020.3. Krok sloužil jako kontrola.

V obou případech byly použity obdobné parametry – pro výpočet polárních solvatačních energií bylo využito dielektrických konstant $\epsilon_{\text{interní}} = 2$, $\epsilon_{\text{solvent}} = 80$, $\epsilon_{\text{vakuum}} = 1$. Výpočet nepolárních solvatačních energií byl proveden pomocí metody SASA.

Na základě výsledků bylo vybráno vždy několik nejlepších modelů. Na vybraných modelech byla pomocí MD trajektorií provedena analýza pomocí aplikace gRINN, ze které byly získány matice intermolekulových interakcí. Pomocí vlastní Python knihovny byly vypočteny “diferenční matice” – procentuální shody síly a pozice vazebných interakcí mezi dvěma maticemi intermolekulových interakcí, dle kterých bylo zjišťováno, jestli nejsou některé modely v rámci jednoho proteinu redundantní. Seznam vybraných modelů je uveden v **tabulce 5**.

Tabulka 5 – Modely získané dokováním serverem ClusPro2, vybrané pro další analýzy na základě MM/PBSA analýz vazebné energie v komplexu s 10E8 protilátkovým fragmentem.

Var.	MLA024	MLA025	MLA092	MLA158	MyoWT	MPER
Modely	2, 3, 5, 6	2, 5, 8	0, 1, 8	8	2, 6, 8	2, 6

4.2.3.4 Analýza vazby, vazebné afinity a termodynamických parametrů

Na vybraných modelech byla provedena „steered MD“ s následnou analýzou nerovnovážné trajektorie pomocí metody Umbrella Samplingu. Vstupními strukturami pro simulace byly modely vybrané pomocí MM/PBSA s použitím gmx_MMPBSA balíku. Tyto struktury byly reprezentovány s „coarse-grained“ rozlišením za využití silového pole SIRAH. Popsány zde budou pouze specifika přípravy systému, nikoliv samotný proces, který je uspokojivě popsán v publikovaném tutoriálu silového pole a v souvisejícím článku.^{93,94}

Proteiny byly vloženy do kubických boxů se vzdáleností nejméně 2.0 nm od hrany a délkou vektoru Z 17 nm. Molekuly komplexů byly orientovány podél principiální osy a v případě potřeby rotovány tak, aby osa vazby 10E8fr-MLAxxx/MyoWT/MPER směřovala ve směru osy Z. Do roztoku byly přidány ionty Na⁺ a Cl⁻ do neutralizace proteinu a následně do celkové koncentrace 0.150 M. Správnost přípravy a orientace struktury byla vizuálně kontrolována pomocí VMD. Po samotné přípravě (solvatace, minimalizace, ekvilibrace) proběhla nerovnovážná produkční MD simulace:

- 250000 kroků, $\delta t = 20 \text{ fs}$, celkově 500 ns, PBC ve všech dimenzích
- Tlakový coupling s Parrinello-Rahman barostatem (parametry identické s těmi popsánymi **kapitolou 4.2.3.1**)
- Tahání MLA/MyoWT/MPER ve směru osy Z s „umbrella“ potenciálem, rychlostí $10 \text{ nm} \cdot \text{ns}^{-1}$ a silou $1000 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$
- Analýza počtu kontaktů mezi proteiny během trajektorie tahání

Výsledná trajektorie byla rozdělena na jednotlivé snímky a pro každý snímek byla vypočtena vzdálenost mezi 10E8fr. a vazebným partnerem. Byly vybrány snímky

odpovídající aditivnímu zvyšování vzdálenosti COM-COM proteinů („center-of-mass-center-of-mass“) o 0.150 nm počínaje nulovým snímkem.

Pro každý z takto vybraných snímků byl proveden samotný “umbrella sampling”. Pro tyto kroky byla snížena rychlost tahání na 0.0 – efektivně byly tedy z hlediska vzdáleností těžišť zamrazeny na místě. Každý snímek byl ekvilibrován po dobu 1 ns v NPT souboru s využitím identických parametrů jako v předchozích krocích. Následovala produkční simulace trvající pro každý snímek 10 ns během kterých byly sbírány síly působící na pomyslnou pružinu udržující vzdálenost mezi těžišti molekul.

Pomocí struktur takto simulovaných snímků trajektorie a průběhů sil působících na pružinu byly s použitím algoritmu WHAM a statistického vyhodnocení chyb metodou b-hist se 100 opakováními v 10 blocích vypočteny histogramy pokrytí trajektorie tahání datovými body a průběhy funkcí PMF podél trajektorie. Výjimku činila vždy první nanosekunda každé trajektorie. Pro získání $\Delta G_{\text{tahání}}$ byly průběhy do vzdálenosti COM-COM proteinů 5 nm analyzovány vlastní Python knihovnou identifikující minimum průběhu odpovídající navázanému stavu s odlišením minima od nízké hodnoty způsobené nedostatečným vzorkováním a maxima. Na základě těchto potenciálů, aproximujících vazebnou energii protein v coarse-grained reprezentaci, byl vybrán vždy jeden nejlepší model pro každého vazebného partnera.

Obdobné kroky byly zopakovány pro tři vybrané modely 10E8fr.-MLA158 komplexů (modely 2,3 a 8) s atomistickou reprezentací pomocí silového pole CHARMM36m. Příprava systémů pro produkční simulaci proběhla v boxech připravených stejným způsobem jako pro coarse-grained simulace, minimalizace a ekvilibrace proběhla ekvivalentně dříve popsaným atomistickým MD simulacím. Tahání, US a WHAM proběhlo s identickými parametry jako při coarse-grained simulacích.

4.2.3.5 IEM

Z trajektorií využitých pro MM/PBSA analýzu byly vybrány ty odpovídající modelům vybraným v předchozím kroku. Přidána byla trajektorie vyřešené struktury 10E8fr.-MPER, a modelů 6 a 8 stejného komplexu. Trajektorie byly upraveny tak, aby byly odstraněny vlivy PBC a zkráceny o prvních 15 ns. Takto získaných 5 ns trajektorie (500 snímků) bylo pro každý snímek analyzováno pomocí aplikace gRINN využívající balík GROMACS se silovým polem CHARMM36m. Výsledné datasety byly analyzovány pomocí vlastní Python knihovny.

Jako další kontrola dokování a MM/PBSA výběru bylo provedeno porovnání komplexů 10E8fr-MPER vycházejících jak z vyřešené struktury komplexu tak z modelů komplexu 6 a 8. Tyto struktury byly společně s ostatními analyzovány ze dvou hledisek, jednak z pohledu celkových a energeticky nejlepších interakcí, tak z pohledu konkrétních interakcí cílících na experimentálně stanovená důležitá rezidua protilátky 10E8.

4.2.3.6 Kontrola postupu pomocí vyřešené struktury 10E8 Fab-MPER komplexu

Vybrané modely komplexů 10E8fr. – MPER získané kombinací dokování a MM/PBSA byly porovnány s analogickou analýzou provedenou na krystalografické vyřešené struktuře stejného komplexu. Byly porovnány konformace nejdříve z geometrického hlediska (RMSD hlavních řetězců ligandů a jejich pozice) pomocí programu UCSF Chimera. Následně bylo porovnání provedeno i z hlediska párových interakcí s 10E8fr.

4.2.3.7 Modelování tetramerního MLA024 proteinu

Na základě prvotního homologního modelu proteinu MLA024 bylo provedeno modelování tetramerní formy MLA024-tetramer. Výsledná struktura byla vizualizována pomocí PyMol a UCSF Chimera, pomocí kterého byly vypočteny možné vodíkové vazby mezi jednotlivými monomery komplexu. Pomocí aplikace HydroPRO byl vypočten translační difuzní koeficient této struktury tetrameru a struktury bakteriální Hsp70 (pdb: 2KHO) a predikce struktury získané z AlphaFold databáze.^{87,90,91}

K výpočtům hydrodynamických vlastností bylo využito odhadu reciproké hustoty proteinu $0,73 \text{ ml} \cdot \text{g}^{-1}$, hustoty solventu $\rho = 1,00 \text{ g} \cdot \text{cm}^{-3}$ a teploty $25 \text{ }^\circ\text{C}$.⁹⁵

Struktura byla simulována MD simulací po dobu 20 ns pod silovým polem CHARMM36m.

5 Výsledky a diskuze

5.1 Selekcce proteinových variant vážících 10E8 IgG a jejich charakterizace

5.1.1 Úvod

Celý postup výzkumného projektu začal vysoce komplexní knihovnou oligonukleotidů DNA, které nesly genetickou informaci pro Myomedinové varianty s variabilní aminokyselinovou sekvencí na 12 pozicích. Tyto proteiny byly produkovány *in vitro* na ribosomech se zablokovanou schopností odpojení polypeptidového řetězce, tedy v komplexu protein-ribosom-mRNA. Metodou ribosomálního displeje s pozitivní selekcí na vazbu protilátky 10E8 a negativní selekcí na isotypovou kontrolu bylo vybráno množství potenciálních vazebných proteinů, jejichž kódující mRNA byla reverzní transkripcí přepsána na DNA. Tato DNA pomocí metody PCR vložena do bakteriálních plasmidů v konstruktech His-MLA-Avi, tímto byla vytvořena výsledná knihovna selektovaných variant v plasmidu.

Tato plasmidová knihovna byla použita k produkci proteinových variant, z nichž byly pomocí ELISA dále selektovány ty varianty, které vykazují preferenční vazbu na 10E8. Získané bakteriální lyzáty byly podrobeny analýze vazebných afinit na terčovou protilátku 10E8, na isotypovou kontrolní protilátku IgG λ a na protein BSA. Na základě získaných výsledků byly vybrány varianty s výrazně vyšší vazebnou afinitou k terčové protilátce oproti kontrolám. Sekvence DNA kódující tyto vybrané varianty byly zjištěny pomocí Sangerova sekvenování a tyto Myomedinové varianty postoupily do dalších kol selekce. Podrobnější stanovení vazebné afinity metodou ELISA, tentokrát již s využitím purifikovaných proteinů, ředících řad a stanovením vazebných křivek umožnilo proteiny seřadit na základě poměru zjištěných hodnot K_d k oběma protilátkám. Nejlepší varianty byly molekulárně biologickými metodami převedeny do konstruktů s jiným tagem na C-konci, Avi Tag byla nahrazen V5-tagem a byla opět stanovena jejich vazebná afinita a specifita. Nejlepších několik variant bylo produkováno a na spolupracujícím pracovišti Prof. Mgr. MUDr. Milana Rašky, Ph.D. (Lékařská fakulta, UPOL) byla provedena série imunizačních experimentů na myším modelu, získaná séra byla poměrně

podrobně charakterizována na přítomnost neutralizujících protilátek schopných blokovat vstup viru do buňky. Experimenty na myším modelu prokázaly významnou schopnost řady variant vybudit produkci neutralizujících sér proti reprezentativnímu panelu pseudovirů HIV-1. Výsledky této práce byly publikovány.⁵¹

Variety prokazující nejpříznivější biologickou aktivitu byly vybrány pro další charakterizace, purifikace SEC i pro *in silico* experimentální část této práce.

5.1.2 Charakter experimentální práce

Charakter experimentální části této práce byl veden několika faktory, z nichž některé budou pečlivěji rozebírány v dalším textu. Proteiny pěstované a charakterizované v této práci vycházely z kombinatoriální knihovny, vybudované na bázi proteinového scaffoldu Myomedin, který sestává ze 111 aminokyselin z nichž 12 je zaměnitelných. Po nezbytných molekulárně biologických a biochemických krocích (selekcce variant z většího množství, záměna terminálních tagů, exprese a imunochemická analýza metodou ELISA) byly varianty základně charakterizovány některými biofyzikálními metodami.

Hlubší biofyzikální analýza proteinů z hlediska struktury (vyřešení struktur komplexů) a vlastností (měření vazebné afinity v roztoku) se ukázala být neuskutečnitelnou, primárně vzhledem k nízké výtěžnosti produkci, problémům s čistotou proteinů a se stabilitou některých z nich. Vazebný partner proteinů, protilátka 10E8, je velmi vzácná a potřebná k provedení řady experimentů. Problematická byla v mnoha chvílích i možnost samotného provádění experimentů, omezená současnou pandemickou situací.

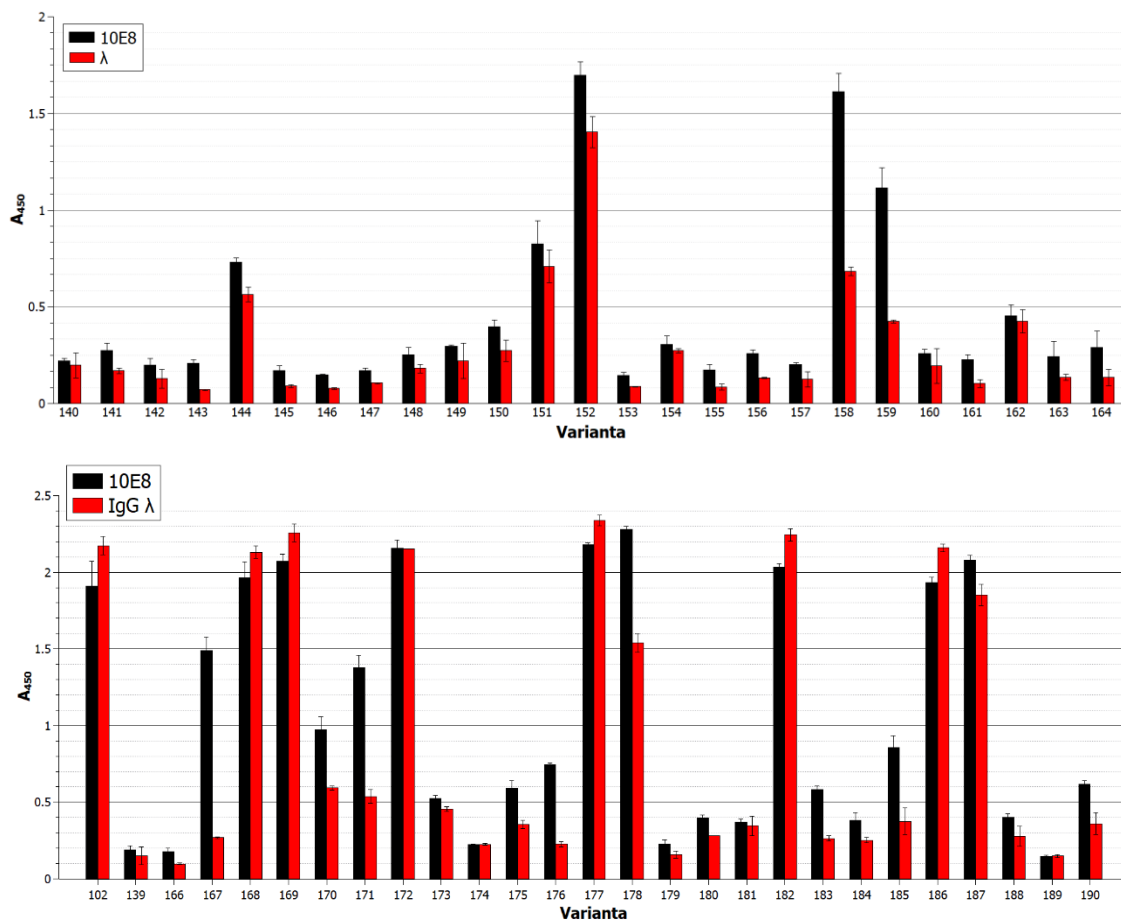
Alespoň část těchto omezení se podařilo obejít pomocí teoretických *in silico* metod, které činí druhou podstatnou část této práce. Nyní budou diskutovány výsledky *in vitro* charakteru primárně molekulárně biologického, biochemického a omezeně biofyzikálního.

5.1.3 ELISA, plošná selekce z proteinové knihovny

Z kombinatoriální knihovny byly metodou ribosomálního displeje selektovány proteinové varianty (a jejich korespondující mRNA) vážící se na 10E8 protilátku. Po reverzní transkripci na cDNA byly tyto sekvence metodou PCR amplifikovány a posléze vloženy do pET28b plasmidů, kterými byly transformovány bakteriální buňky XL1-Blue. Byla odebrána a kultivována řada kolonií, z jejichž kultur byly izolovány plasmidové DNA (tento postup nebyl předmětem práce autora a je uváděn pro kontext).

Tyto plasmidové DNA byly vloženy do bakteriálních buněk BL21 (DE3) pBirA optimalizovaných pro expresi rekombinantních proteinů s biotinylací Avi-tagů přímo v buňkách. Narostlé kolonie byly kultivovány, lyzovány sonikací a přímo na lyzátech bylo provedeno stanovení poměru vazebné afinity produkovaných klonů na 10E8 oproti kontrolní IgG λ metodou ELISA. Experiment byl prováděn v triplicátech a s negativní kontrolou BSA. Koncentrace proteinu v lyzátu nebyla stanovena, proto odhad síly vazby byl pouze orientační, avšak pro daný účel výběru variant plně dostačující. Touto metodou bylo charakterizováno celkem téměř 200 klonů.

Výsledky jsou formou sloupcových grafů vizualizovány na **obrázku 10**.



Obrázek 10 – Stanovení poměru vazby His-MLA-Avi variant na 10E8 a isotypovou kontrolní protilátku IgG λ pomocí ELISA. Experiment byl prováděn v triplicátech a chybové úsečky odpovídají standardní odchylce.

Na základě zde zveřejněných výsledků i dalších byly vybrány klony s příznivými poměry vazeb. Ty byly ověřeny sekvenováním a na základě toho se výběr klonů zúžil na varianty MLA014, 015, 016, 024, 025, 045, 072, 092, 093, 132, 158 a 159. Některé klony exprimovaly zkrácený protein v důsledku mutací měnící aminokyseliny na předčasný stop kodon nebo inserce či delece bází posunuly rámec čtení při translaci tak, že za tímto místem následovala alternativní čtená sekvence ukončená předčasným STOP kodonem. Například klon MLA167, který ač ukazuje příznivé vazebné signály v ELISA, obsahuje posun čtecího rámce, avšak vzniká funkční Avi Tag, proto je protein v ELISA detekován (Kuchař, ústní sdělení). Tyto klony byly z další charakterizace vyloučeny.

Po sérii experimentů, která prokázala, že je C-terminální Avi-tag nevhodný z hlediska specifity vazby (nepublikované výsledky) byly všechny varianty převedeny do konstruktů His-MLA-V5, které jsou předmětem zbytku této práce.

5.1.4 Záměna C-terminálního Avi-tagu za V5-tag

Z namnožených plasmidů kódujících varianty v konstruktech His-MLA-Avi byla pomocí štěpení restrikasami NcoI a BamHI-HF vyštěpena část kodující příslušnou variantu Myomedinu MLA (NcoI-His-MLA-BamHI).

Rozdělené sekvence DNA byly separovány pomocí elektroforézy v agarosovém gelu v TAE pufru. Inerty byly vyříznuty skalpelem a izolovány pomocí QIAquick Gel Extraction Kitu. Takto získané úseky DNA byly ligovány do pET28b plasmidu, kde na C-konci expresní kazety byla sekvence kódující V5-tag (pET-28b-NcoI-BamHI-V5 Tag). Byla provedena transformace buněk XL1-Blue včetně jedné kontrolní (s ligací směsí bez obsahu insertu). Výsledkem byla jediná bakteriální kolonie na agarové misce pro kontrolní ligaci a řádově stovky kolonií na ostatních miskách.

Byly izolovány plazmidy pomocí QIAprep Spin Miniprep Kitu a transformovány do expresních buněk BL21 (DE3).

5.1.5 Produkce, purifikace proteinů pomocí Ni-NTA agarosy pro analýzu vazebné afinity

Proběhlo napěstování bakteriální kultury, indukce pomocí IPTG, exprese proteinů, lýze buněk sonikací a purifikace na Ni-NTA agarose s využitím N-terminálního His-tagu.

Byla měřena absorbance získaných roztoků při $\lambda = 280$ nm. Na základě predikovaných extinkčních koeficientů získaných pomocí webového serveru ExpASY ProtParam server byly vypočteny hmotnostní koncentrace proteinů. Na základě přibližné molekulové hmotnosti každého proteinu 16 000 Da byly vypočteny molární koncentrace. Výsledky jsou uvedeny v **tabulce 6**. Je třeba zmínit, že jak tyto výsledky, tak výsledky stanovení afinit jsou přibližné, extinkční koeficienty jsou predikované, purifikované frakce navíc obsahují i určitý podíl bakteriálních proteinů.

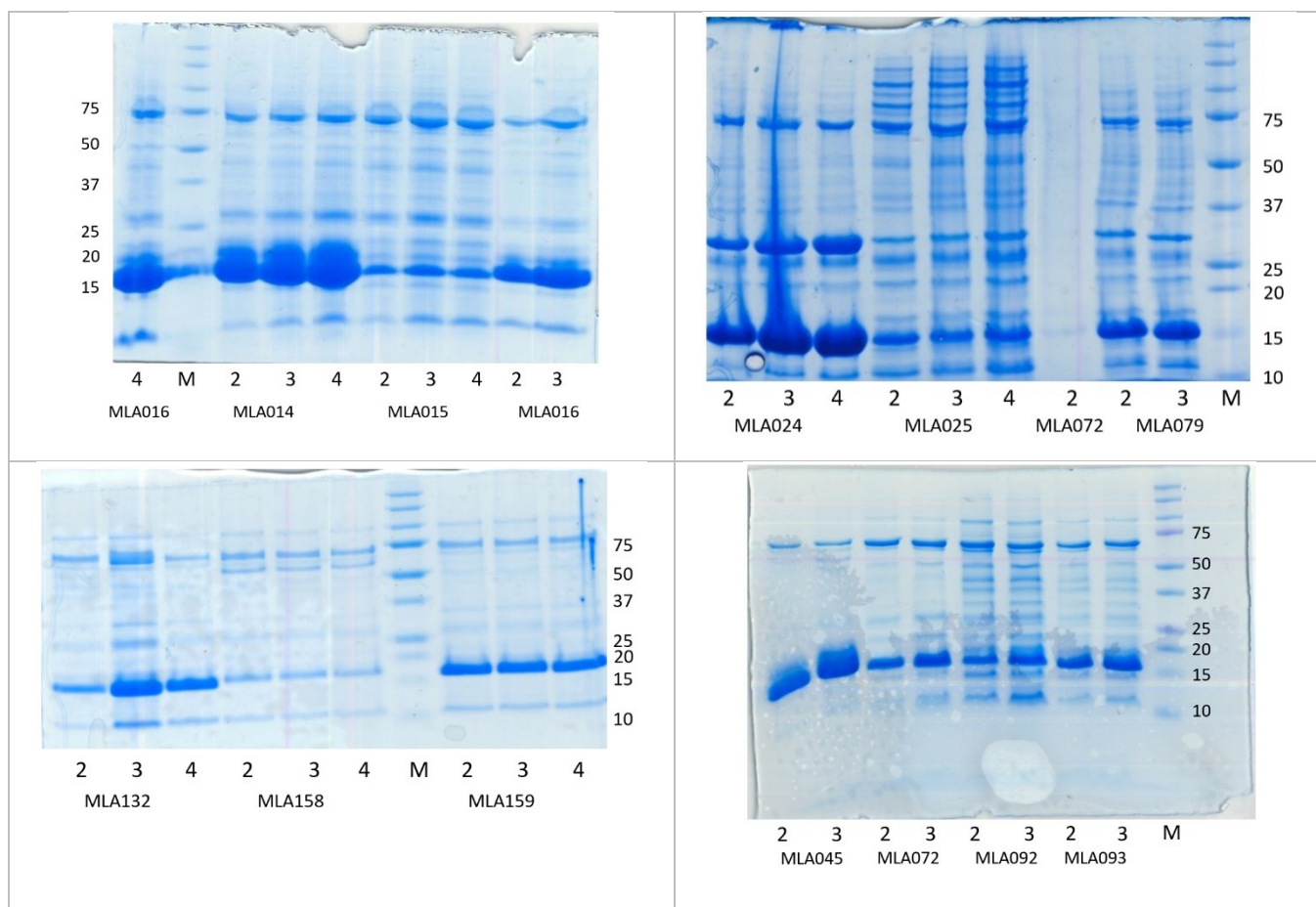
Tabulka 6 – Predikované extinkční koeficienty (v podobě abs. ekvivalentů), naměřené A_{280} a z nich vycházející koncentrace třetích frakcí proteinů získaných purifikací na Ni-NTA agarose.

His-MLA-V5 (fr. 3) var.	Abs. ekv. $c = 1 \text{ g} \cdot \text{dm}^{-3}$	A_{280}	$c_m / \text{mg} \cdot \text{ml}^{-1}$	$c / \text{mol} \cdot \text{dm}^{-3}$
MLA014	0,943	5,31	5,63	$3,52 \cdot 10^{-4}$
MLA015	1,033	1,00	0,968	$6,05 \cdot 10^{-5}$
MLA016	0,945	3,80	4,02	$2,51 \cdot 10^{-4}$
MLA024	0,940	0,42	0,447	$2,79 \cdot 10^{-5}$
MLA025	0,954	0,14	0,147	$9,19 \cdot 10^{-6}$
MLA045	0,950	1,09	1,15	$7,19 \cdot 10^{-5}$
MLA072	0,950	0,27	0,28	$1,75 \cdot 10^{-5}$
MLA079	1,29	0,66	0,510	$3,19 \cdot 10^{-5}$
MLA092	1,65	0,83	0,503	$3,13 \cdot 10^{-5}$
MLA093	1,30	0,70	0,70	$3,38 \cdot 10^{-5}$
MLA132	1,72	1,10	0,639	$3,99 \cdot 10^{-5}$
MLA158	2,47	0,15	0,061	$3,81 \cdot 10^{-6}$
MLA159	1,09	0,25	0,229	$1,43 \cdot 10^{-5}$

Již na těchto výsledcích lze pomocí koncentrací lišících se v rozsahu dvou řádů, že i přes identické podmínky proteinové produkce jsou výtěžky po purifikacích pomocí Ni-NTA agarosy značně odlišné mezi variantami. Nejhoršího výtěžku bylo z uvedených variant získáno u MLA025 a MLA158, nejlepšího u MLA014 a MLA016.

Výsledky SDS-PAGE elektroforéz těchto variant dále ilustrující tyto rozdíly jsou zobrazeny v **tabulce 7**. Pozice MLA proteinů na SDS-PAGE gelech leží těsně nad pruhem proteinového markeru označujícího 15 kDa (MLA proteiny mají velikost přibližně 16 kDa) – tyto pruhy jsou u řady variant velice výrazné, ilustrující dobrou produkci. Kromě očekávaných pruhů okolo 16 kDa jsou patrné i všudypřítomné, nezanedbatelné pruhy okolo 75 kDa a v mnohých případech (zejména MLA024) i výrazné pruhy nad 25 kDa bandem markeru.

Tabulka 7 – Barvené SDS-PAGE gely po elektroforéze proteinových směsí purifikovaných pomocí Ni-NTA. Písmeno M označuje proteinový marker s vyznačenými velikosti proteinů v kDa, čísla 2-4 označují eluovanou frakci z Ni-NTA purifikace, pod každým gelem je uvedena varianta odpovídající konkrétním pruhům.



5.1.6 Analýza vazebné afinity metodou ELISA

Afinita k protilátkám 10E8 a IgG λ byla u proteinových variant MLA (His-MLA-V5) purifikovaných na Ni-NTA agarosové matrici stanovena pomocí fitování křivek získaných pomocí ELISA. Vzhledem k jen přibližné znalosti koncentrací nanášených vazebných proteinů je nutné výsledky brát pouze orientačně. Uspokojivě lze mluvit o řádech (varianty vážou 10E8 protilátku s afinitami K_d v řádu jednotek až desítek nanomolů na litr), nikoliv o přesných hodnotách. Jak bude patrné v příštích kapitolách, nejsou zároveň Ni-NTA purifikované proteiny dostatečně čisté.

Bylo provedeno i porovnání vazebných afinit MyoWT ve verzi s V5 i Avi-tagem. Shrnutí stanovených hodnot K_d je uvedeno v **tabulce 8**. Stanovení hodnot K_d proběhlo proložením sigmoidní funkcí. Tabulka obsahuje některá nepřesná stanovení hodnot, což bylo způsobeno tím, že byl hledaný inflexní bod sigmoidy (K_d) mimo stanovovaný rozsah koncentrací. Tato skutečnost je u vazeb na IgG λ příznivá, neboť to znamená, že K_d je u IgG λ vyšší, než je maximum analyzovaného rozsahu (řádově μM koncentrace) a MLA protein tedy váže tuto kontrolní protilátku mnohem slabší afinitou. V takovém případě byly hodnoty, pokud to bylo možné, odhadnuty a jsou nepřesné.

Tabulka 8 – Disociační konstanty (K_d) stanovené pomocí ELISA. Touto metodou byly stanoveny K_d vazeb MLA proteinů s 10E8 a IgG λ protilátkami. Uvedené střední hodnoty K_d byly získány proložení sigmoidní funkcí. Varianty jsou seřazeny dle poměru středních hodnot K_d 10E8 a IgG λ, Větší poměr naznačuje významnější preferenci vazby k 10E8, hvězdička (*) značí nepřesný odhad hodnoty.

His-MLA-V5	K _d (10E8) / nM	K _d (IgG λ) / nM	$\frac{K_d(\text{IgG } \lambda)}{K_d(10\text{E8})}$
MLA158	10,8	1000*	92,6*
MLA014	11,7	921	78,7
MLA015	3,11	241	77,7
MLA132	31,9	*	*
MLA159	45,8	1000*	21,8*
MLA016	7,32	86,7	11,9
MLA093	106	1000*	9,43*
MLA092	21,1	183	8,67
MLA025	22,6	152	6,73
MLA079	33,6	164	4,9
MLA024	300*	1000*	3,3*
MLA072	91,5	229	2,50
MLA045	*	*	*
<i>MyoWT (V5)</i>	274	6530*	23,8*
<i>MyoWT (Avi)</i>	6,8*	17,7*	2,60*

Získané křivky včetně proložení jsou dostupné v **přílohách A 1 až A 5**.

Výsledky ukazují, že v rámci His-MLA-V5 variant váže většina stanovovaných proteinů 10E8 protilátku výrazně lépe, než His-MyoWT-V5 protein, který zároveň špatně váže IgG λ. Jak bylo zmíněno dříve, ukázaly se C-terminální Avi-tagy jako nevhodné. Toto je ilustrováno i vazebnou afinitou His-MyoWT-Avi, která je na 10E8 protilátku v řádech jednotek nM. Stejně tak existuje i jen o řád slabší vazba na IgG λ. Toto je konzistentní s pozorovaným zlepšením specifity (reprezentovaného rozdíly vazby na obě protilátky) které nastalo po záměně Avi-tagů za V5-tagy (výsledky nepublikovány).

Stále však existuje nezanedbatelná vazba His-MyoWT-V5 na 10E8 protilátku. Vazbou samotného MyoWT (bez tagů) se blíže zabývá *in silico* část této práce (**kap. 5.2.9.1.7**). Nelze přitom jednoznačně říct, že by byly výsledky z ELISA přenositelné do *in silico* výpočtů, vzhledem k charakteru imobilizace proteinů, jejich nedefinované čistoty a

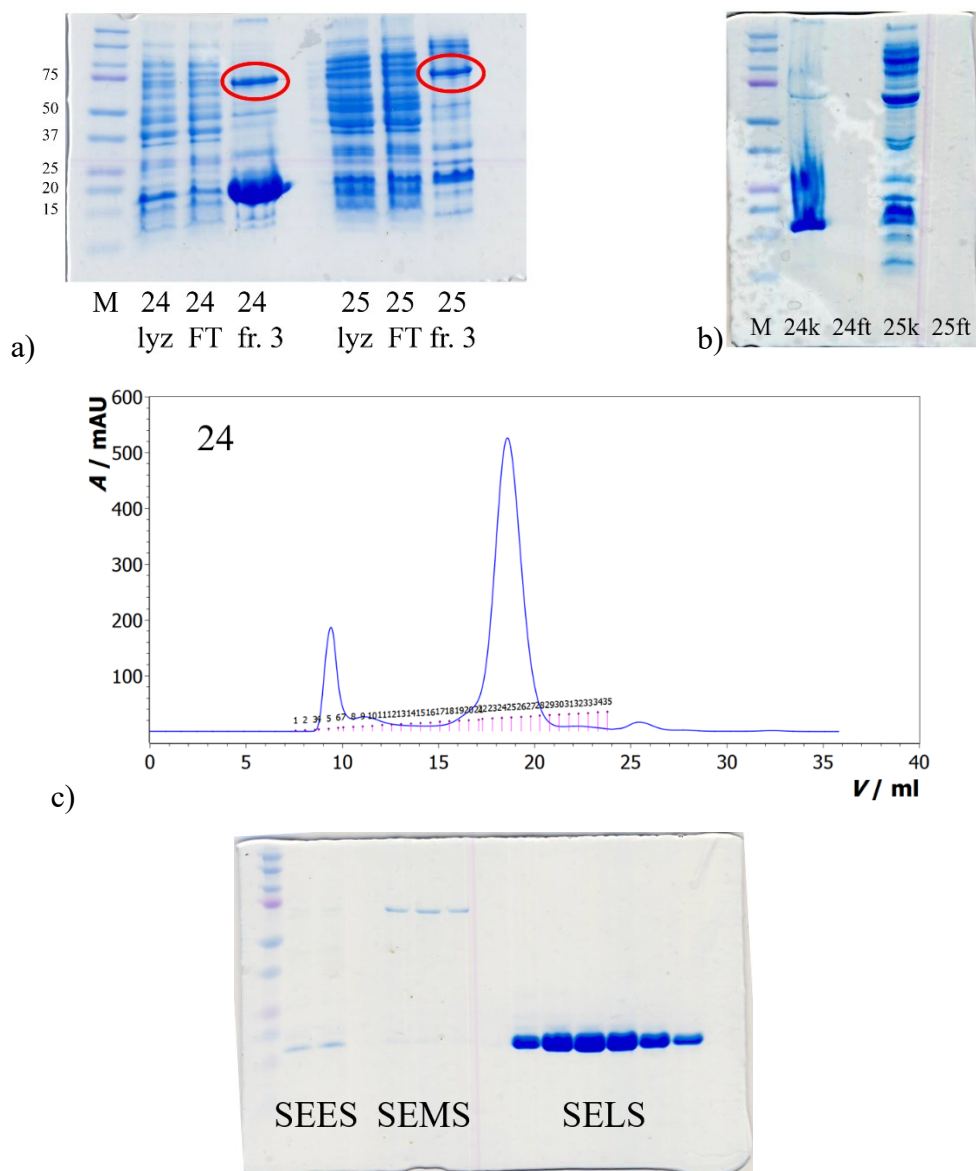
detekce pomocí tagů, která může přinášet bias stanovení. Stanovení na pevném nosiči desky může ve skutečnosti analyzovat jiné vazebné módy, než jaké by existovaly v roztoku (a jsou predikovány v *in silico* části). Zdá se však, že jsou *in silico* výsledky v souhlasu se zde prezentovanými experimentálními (viz **kapitola 5.3**).

5.1.7 Purifikace SEC

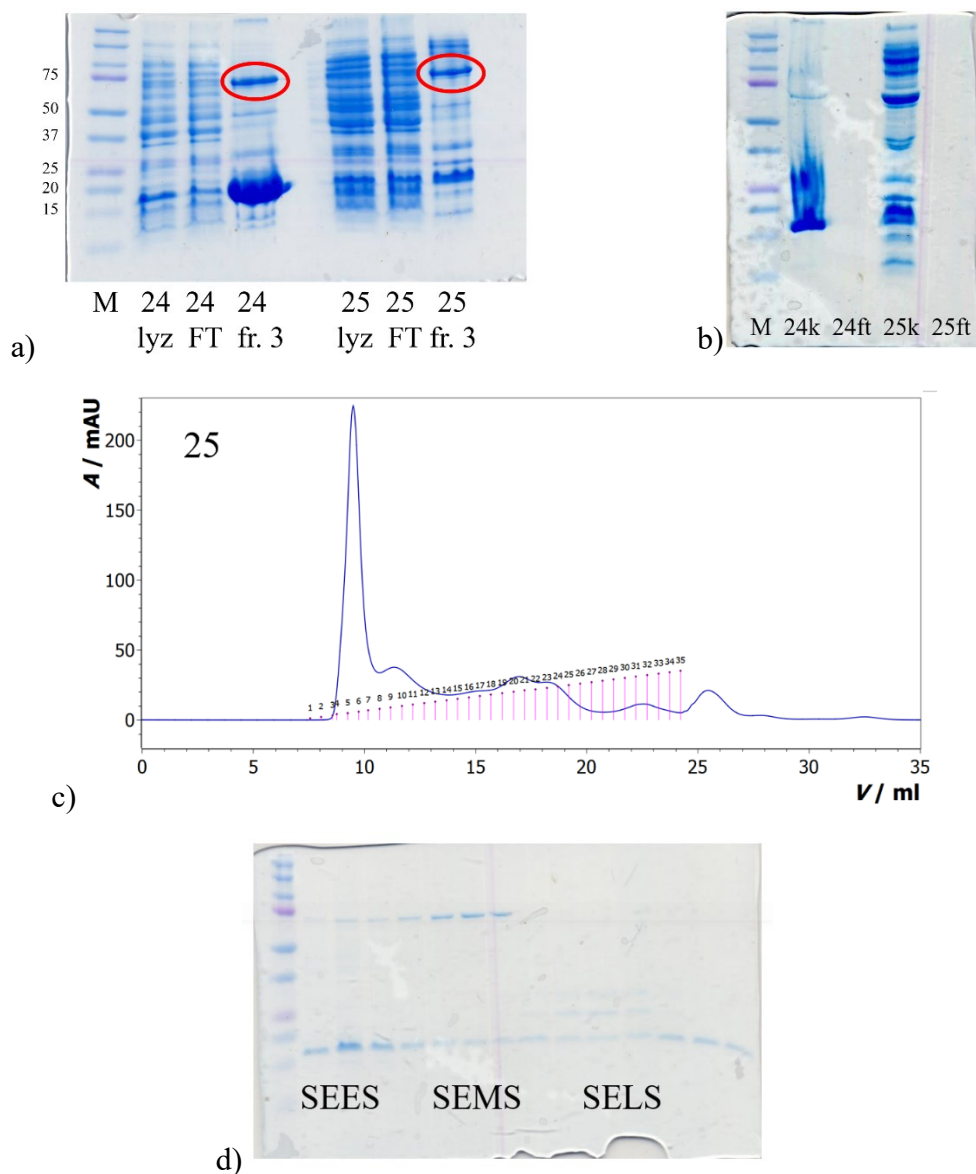
Před provedením této části práce začaly být známy výsledky vakcinačních experimentů. Na základě nich byla vybrána pro další charakterizaci varianta MLA158, MLA092, MLA025. Dále byla vybrána varianta MLA024 jakožto nejlépe se produkující varianta z těch, které alespoň s malou, ale prokazatelnou měrou navozují tvorbu neutralizujících protilátek při imunizaci myší. Pro porovnání byla zároveň provedena i charakterizace MyoWT.

Varianty MLA024 a MLA025 byly pěstovány v 500 ml média, MLA092 a MLA158 ve 2400 ml média, MyoWT ve 200 ml. U varianty MLA158 byla provedena série pokusů optimalizovat podmínky s cílem navýšit výtěžky proteinu, zahrnující optimalizace média, pufrů a teplotních podmínek. Proteinové varianty byly purifikovány pomocí Ni-NTA agarosy, zakoncentrovány vždy do 0,5 ml a poté purifikovány gelovou permeační chromatografií na základě své molekulové hmotnosti (hydrodynamického poloměru), přičemž byly jímány frakce dle absorbancí při $\lambda = 280$ nm. Frakce odpovídající jednotlivým píkům byly sloučeny a podle potřeby zakoncentrovány pro další analýzy. Byly provedeny SDS-PAGE elektroforézy a stanovení koncentrací frakcí.

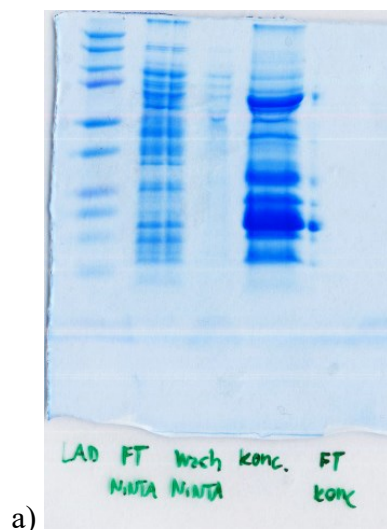
Výsledné SDS-PAGE gely (pokud byly provedeny) a chromatogramy jsou zobrazeny na **obrázcích 11 až 15**. Tři pro každou purifikaci charakteristické píky byly pojmenovány podle pořadí eluce – SEES (size-exclusion early series), SEMS (size-exclusion middle series), SELS (size-exclusion late series).



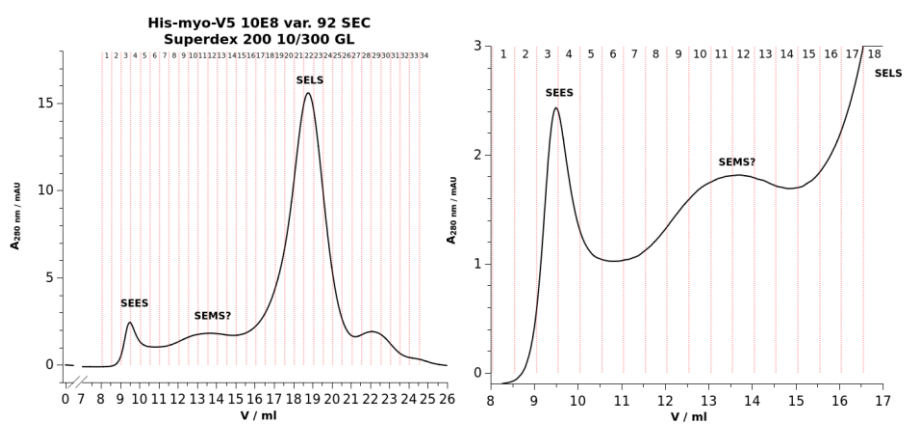
Obrázek 11 – Výsledky SEC purifikace His-MLA024-V5. a) SDS-PAGE barvený gel frakcí z Ni-NTA purifikace (lyz = lyzát, FT = průtok po nanášení kolony, fr. 3 = třetí eluovaná frakce) b) SDS-PAGE barvený gel po zakoncentrování Ni-NTA purifikovaných proteinů (k = koncentrovaný roztok, ft = průtok filtrem koncentrátoru) c) Chromatogram (závislost absorbance při $\lambda = 280$ nm na protékajícím objemu mobilní fáze) d) SDS-PAGE barvený gel jednotlivých frakcí z SEC (SEES = první pík, SEMS = druhý pík, SELS = třetí pík/největší pík). Červené kroužky označují výrazné pruhy o velikosti mírně nižší, než 75 kDa.



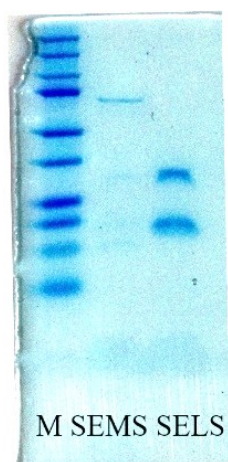
Obrázek 12 – Výsledky SEC purifikace His-MLA025-V5. a) SDS-PAGE barvený gel frakcí z Ni-NTA purifikace (lyz = lyzát, FT = průtok po nanášení kolony, fr. 3 = třetí eluovaná frakce) b) SDS-PAGE barvený gel po zakoncentrování Ni-NTA purifikovaných proteinů (k = koncentrovaný roztok, ft = průtok filtrem koncentrátoru) c) Chromatogram (závislost absorbance při $\lambda = 280 \text{ nm}$ na protékajícím objemu mobilní fáze) d) SDS-PAGE barvený gel jednotlivých frakcí z SEC (SEES = první pík, SEMS = druhý pík, SELS = třetí pík/největší pík). Červené kroužky označují výrazné pruhy o velikosti mírně nižší, než 75 kDa.



a)

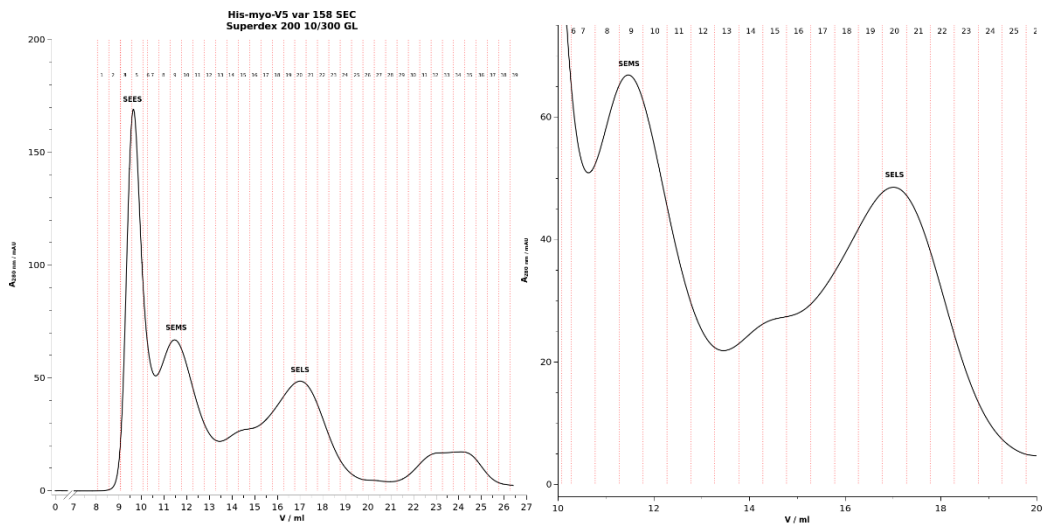


b)

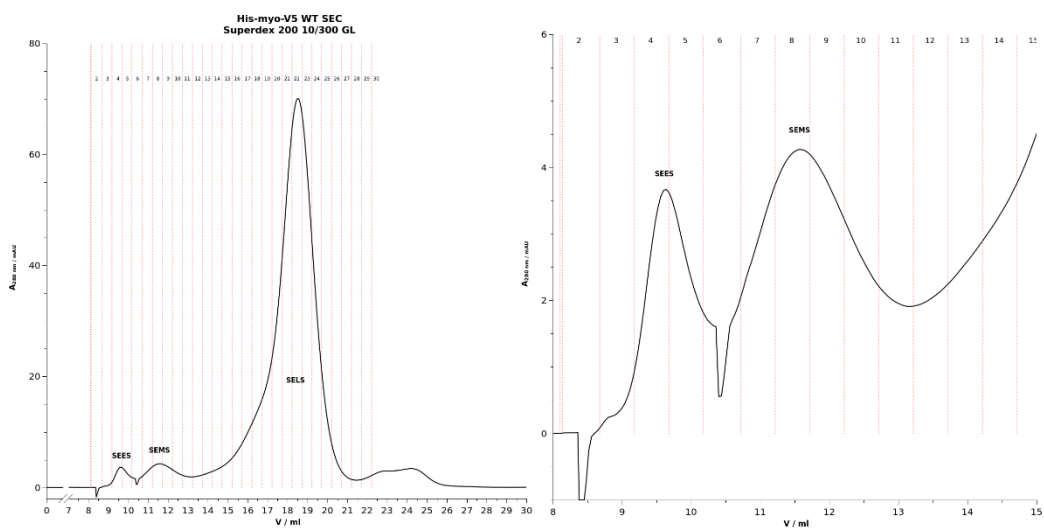


c)

Obrázek 13 – Výsledky SEC purifikace His-MLA092-V5. a) SDS-PAGE barvený gel frakcí z Ni-NTA purifikace (lyz = lyzát, FT = průtok po nanášení kolony, fr. 3 = třetí eluovaná frakce) b) Chromatogram (závislost absorbance při $\lambda = 280 \text{ nm}$ na protékajícím objemu mobilní fáze) c) SDS-PAGE barvený gel jednotlivých frakcí z SEC (SEES = první pík, SEMS = druhý pík, SELS = třetí pík/největší pík)



Obrázek 14 – Chromatogram SEC purifikace His-MLA158-V5 (závislost absorbance při $\lambda = 280$ nm na protékajícím objemu mobilní fáze).



Obrázek 15 – Chromatogram SEC purifikace His-MyoWT-V5 (závislost absorbance při $\lambda = 280$ nm na protékajícím objemu mobilní fáze).

Sumarizace těchto experimentů je následující:

- a) Jako první stojí za povšimnutí SDS-PAGE gely provedené po Ni-NTA purifikaci. Stejně jako dříve obsahují kromě očekávaného proteinu předpokládané velikosti (16 kDa) i výrazný pruh v pozici pod 75 kDa proteinového markeru. Tato vlastnost je společná všem variantám. Varianta MLA092 má navíc i velmi výrazný pruh na pozici nižší, než 37 kDa (32 kDa by přitom velikostně odpovídal dimeru). Jelikož SDS-PAGE představuje denaturující prostředí, které ruší většinu protein-proteinových interakcí, je velmi pravděpodobné, že tyto pruhy budou odpovídat proteinům o jednom kovalentním řetězci (neočekává se např. slabě nekovalentně interagující dimer) nebo agregátu.
- b) Přítomnost více proteinů se potvrdila na SEC chromatogramech. Tyto experimenty umožňují detekovat existenci nekovalentních interakcí. Ve všech variantách existují tři charakteristické píky (na SDS-PAGE gelech pak 2). Pík frakce SELS MLA092 není zcela symetrický, pravděpodobně zahrnuje více proteinů nebo jeho forem a SDS-PAGE gel obsahuje výrazný pruh na zhruba dvojnásobku molekulové hmotnosti MLA proteinu. Musí tedy existovat protein, který v SEC putuje ve formě komplexu, ale v SDS-PAGE se rozpadá.
- c) Pořadí píků na SEC odpovídá hydrodynamickému poloměru. První vytékající protein by tedy měl být větší, než druhý atd. SDS-PAGE gel variant MLA024 a MLA025 tomuto zdánlivě odporuje. První pík (teoreticky největší protein) na SDS-PAGE putuje jako pruh nápadně připomínající monomerní MLA protein. Naopak druhý pík (teoreticky menší protein než první) putuje jako pruh připomínající dříve pozorovaný pruh velikostně pod 75 kDa pruhem markeru. Zároveň je patrné, že se rozpadá na dva, z nichž jeden putuje stejně jako MLA protein. Třetí pík již putuje jako očekávaný MLA protein (16 kDa). Tento zdánlivý nesoulad je v krátkosti vysvětlen **kapitolou 5.2.11**.
- d) Poměr intenzit píků je značně překvapující. MLA024 má SELS jako majoritní pík, MLA025 jako minoritní. MLA092 jako majoritní, MLA158 jako minoritní. MyoWT má pak z většiny jen SELS pík, ostatní jsou zanedbatelné ale stále existují. Tyto změny mohou záviset na samotných variantách, nebo na podmínkách jejich produkce.

Na základě uvedených skutečností začala být formována hypotéza o existenci komplexů MLA proteinů.

5.1.8 ELISA jednotlivých separovaných SEC frakcí

Pro jednotlivé získané frakce z SEC purifikací a pro proteiny purifikované pouze pomocí chromatografie na Ni-NTA agarose byly provedeny stanovení vazebné afinity obdobné, jaké byly provedeny dříve. Kvůli experimentálním problémům byla tato stanovení provedena pouze pro varianty MLA024, MLA025 a MLA092. Výsledné experimentální body s proložením sigmoidou jsou zobrazeny na **přílohách A 6 až A 8**. Číselně jsou výsledky udány v **tabulce 9**.

Tabulka 9 – Disociační konstanty (K_d) stanovené pomocí ELISA. Uvedené střední hodnoty K_d byly získány proložením sigmoidní funkcí. V závorkách jsou uvedeny hodnoty stanovené na proteinu purifikovaném jen pomocí Ni-NTA. (*) nepřesné stanovení/odhad/nepodařilo se stanovit.

Varianta	Frakce	$K_d(10E8)$ / nM	$K_d(IgG \lambda)$ / nM
His-MLA024-V5	SEES	211	5950*
	SEMS	269	5050
	SELS	5800*	*
	Ni-NTA	211 (300*)	2280 (1000*)
His-MLA025-V5	SEES	151	5130*
	SEMS	256	8980
	SELS	1800*	13000*
	Ni-NTA	126 (22,6)	1490* (152)
His-MLA092-V5	SEES	*	*
	SEMS	*	*
	SELS	195	*
	Ni-NTA	*(21,1)	*(183)

Frakce SELS, které by měly odpovídat monomerním MLA proteinům váží u variant MLA024 a MLA025 dle experimentů s nejnižší afinitou. U těchto variant naopak s vysokou afinitou a specifitou vážou frakce SEMS a SEES, odpovídající větší velikosti než té monomerního myomedinu. U varianty MLA092 se v důsledku nízkému výtěžku produkce a purifikace podařilo stanovit pouze příznivé K_d pro vazbu na 10E8 protilátku.

Při pohledu na SDS-PAGE gely jednotlivých frakcí z SEC purifikace (**viz obrázky 11 až 13**) je zjevné, že se ve všech frakcích vyskytuje protein odpovídající velikosti monomerního Myomedinu.

Kombinací doposud popsaných experimentů bylo prokázáno několik věcí. V nativních podmínkách gelové permeační chromatografie putuje Ni-NTA purifikovaný roztok ve formě tří významných forem. Dvě ze tří těchto forem (SEES, SELS) mají v denaturujících podmínkách SDS-PAGE obdobné chování, putují jako protein o velikosti monomerního myomedinu. Frakce SEMS putuje na SDS-PAGE jako dva pozorovatelné proteiny. Jeden o velikosti monomerního myomedinu, druhý o velikosti nižší, než 75 kDa.

V kombinaci s prezentovanou schopností vazby na 10E8 všech frakcí byla stanovena hypotéza, že samotný myomedin existuje ve třech formách v nativních podmínkách. Jednotlivé formy budou popsány v **kapitole 5.1.9**.

5.1.9 Biofyzikální metody a charakter myomedinových forem

5.1.9.1 Úvodní slovo, tetramerní forma myomedinu

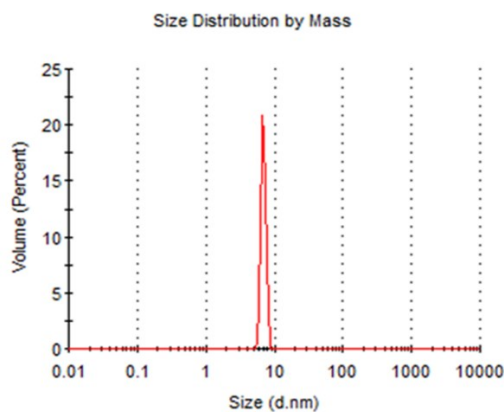
Z hlediska biofyzikálních metod proběhla řada měření na řadě roztoků proteinů MLA. V důsledku nízkého výtěžku rozpustného proteinu, agregace SEC purifikovaných frakcí, experimentálních omezení způsobených současnou pandemickou situací a souvisejících problémů při měření metod není možné na základě výsledků z některých metod (DLS) spolehlivě vyvozovat závěry o vlastnostech analyzovaných proteinů. Ze stejných důvodů nebyly provedeny plánované další metody, které by vnesly jasné důkazy některých hypotéz (SAXS, SPR, MST). Tato kapitola tak bude popisovat pouze malé množství výsledků, které byly ověřeny a jsou spolehlivé. Měření popsána touto kapitolou se nepodařila pro celou řadu možných variant/frakcí provést. Buď nebylo získáno dostatečné množství proteinu, nebo protein silně agregoval.

Velmi diskutovanou hypotézou, která zde nemůže být uspokojivě prověřena, je existence tetramerní formy His-MLA-V5 proteinů, projevující se jako SEES pík na chromatogramech z gelových permeačních chromatografií. Je jisté, že existuje

oligomerní forma myomedinu, která je v nativních podmínkách svým hydrodynamickým poloměrem blízká poloměru Hsp70 proteinu (přibližně 69 kDa).⁹⁶ Tento neznámý oligomer se eluuje dříve, než Hsp70, což by naznačovalo, že by měl být jeho hydrodynamický poloměr větší, nejbližší homo-oligomer myomedinu, pentamer, by měl molekulovou hmotnost přibližně 5x 16 kDa, tedy 80 kDa. Na základě zde neprokázané hypotézy se preferovaným vysvětlením stal myomedinový tetramer (4x 16 kDa, tedy 64 kDa), který by svojí menší velikostí měl eluovat později, než Hsp70. SEES pík se na SDS-PAGE rozpadá na jediný pruh, odpovídající velikostně monomernímu Myomedinu, což naznačuje homo-oligomerní charakter.

Tento hypotetický tetramer byl modelován (viz **Kapitola 5.2.11**) a ve všech získaných modelech existoval v „planární konformaci“. Taková planární konformace by oproti běžné globulární podobě proteinů mohla způsobit pomalejší eluci skrze snížení difuzního koeficientu. Výpočet translačních difuzních koeficientů ve stejné kapitole potvrdil, že by toto bylo principiálně možné.

Jediný částečně uspokojivý *in vitro* experimentální důkaz existence takového tetrameru přineslo DLS měření frakce SEES varianty His-MLA158-V5, které ukázalo naměřený hydrodynamický průměr (maximum píku) $D_H = 7,071$ nm, což by odpovídalo proteinu o velikosti přibližně 64 kDa. Toto je však zatíženo předpokladem globulárního tvaru, který jak bylo diskutováno, nemusí být pro tetramerní myomedin dokonale splněn. Křivka získaná DLS měřením je zobrazena na **obrázku 16**. Analýzy tohoto píku zabudovaným softwarem přístroje Zetasizer Nano ZS90 ukázaly, že je modus $D_H = 7,071 \pm 0,5294$ nm a odhadovaná molekulová hmotnost $M_r = 64,6 \pm 15,1$ nm kde jsou chyby měření uvedeny standardní odchylkou.



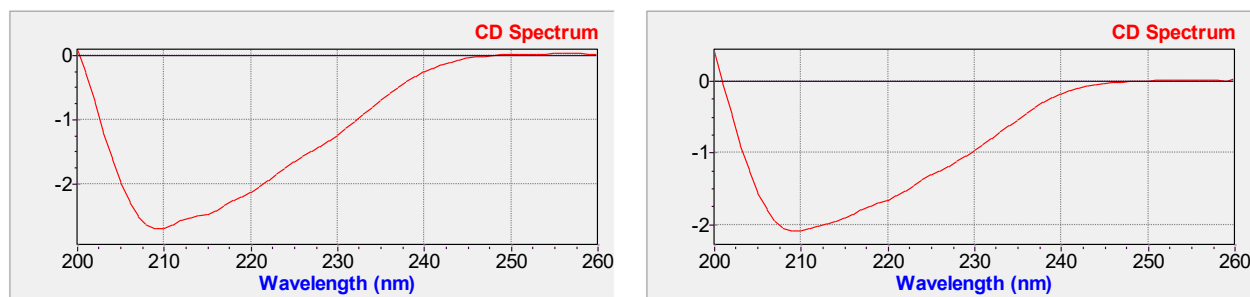
Obrázek 16 – DLS měření frakce SEES His-MLA158-V5. Osa X označuje hydrodynamický průměr částice, osa Y objem zabraný částicí této velikosti v roztoku.

Hypotéza o homo-tetramerním charakteru frakce SEES MLA proteinů je tak stále neprokázána. Důkazy jsou prozatím neuspokojivé a bylo by vhodné je ověřit.

5.1.9.2 Cirkulární dichroismus

Pro biofyzikální charakterizace proteinových frakcí získaných SEC purifikací byly vybrány metody vhodné pro charakterizaci malých množství proteinů. Pro získání informace o sekundární struktuře frakcí SEES a SEMS (ne-monomerní frakce) a SELS (monomerní frakce) His-MLA024-V5 bylo vyzkoušeno měření cirkulárního dichroismu. Kvůli technické chybě se podařilo naměřit pouze CD spektra frakce SELS proteinu His-MLA024-V5 v porovnání s His-MyoWT-V5. Měření bylo provedeno v pufru PBS, MyoWT protein byl purifikován pouze metodou Ni-NTA. Samotný měřicí experiment nebyl proveden autorem (kdo?). Výsledná CD spektra včetně proložením získaného odhadovaného zastoupení sekundárních strukturních elementů jsou viditelná v **tabulce 10**.

Tabulka 10 – Výsledky měření cirkulárního dichroismu frakcí proteinů MyoWT (Ni-NTA) a MLA024 (SELS) v konstruktech s His a V5 tagy. Levá část zobrazuje CD spektrum His-MyoWT-V5 (Ni-NTA purifikovaná frakce), pravá spektrum His-MLA024-V5 (frakce SELS). Ve spodní část jsou uvedeny proložení získané procentuální zastoupení elementů sekundárních struktur.



His-MyoWT-V5 (Ni-NTA)

Helixy	25.1 %
β -Antiparalelní	11.7 %
β -Paralelní	11.2 %
β-otočka	18.7 %
Nestrukt.	37.7 %
Σ	104.3 %

His-MLA024-V5 (SELS)

Helixy	22.0 %
β -Antiparalelní	13.2 %
β -Paralelní	12.6 %
β-otočka	19.4 %
Nestrukt.	41.1 %
Σ	108.2 %

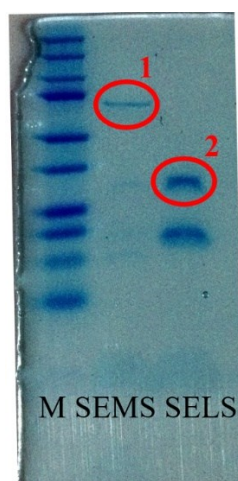
Z výsledků měření CD spektra plyne, že je (v rámci chyby) zastoupení elementů sekundární struktury u SELS frakce His-MLA024-V5 stejné, jako u His-MyoWT-V5. Tato skutečnost podporuje hypotézu, že SELS frakce nejspíše bude monomerní forma myomedinu a také, že aminokyselinovými záměnami nebyla v případě této varianty narušena sekundární struktura. Spektra frakcí SEMS a SEES se nepodařila naměřit a pro analýzy dalších variant bylo od této metody upuštěno. Tento dílčí výsledek však dobře zapadá do celkové mozaiky charakterizace MLA proteinů.

CD spektra byla naměřena pouze v rozmezí vlnových délek 200 nm – 260 nm. Pro spolehlivé stanovení zastoupení elementů sekundární struktury (zejména β-struktur) by bylo třeba získat data z oblasti od alespoň 190 nm. Změření přesného zastoupení elementů sekundární struktury zde však nebylo cílem.

5.1.9.3 Hmotnostní spektrometrie, dimery

Z SDS-PAGE gelu purifikovaných SEC purifikovaných frakcí His-MLA092-V5 byly vyříznuty dva proteinové pruhy a ty byly předány k analýze do *Centra Molekulové Struktury, BIOCEV*. K porovnání byla využita známá sekvence proteinu MLA092 a sekvence bakteriálního chaperonu Hsp70 (DnaK). Sekvence Hsp70 byla vybrána, protože vlastnosti Hsp70 odpovídaly pozorovaným vlastnostem proteinového pruhu vyskytujícího se v SDS-PAGE gelech pod 75 kDa. Hsp70 odpovídá pozorovaným pruhům velikostně, svojí kopurifikací na Ni-NTA agarose a svým rozpadem v denaturujících podmínkách na dva proteiny, z nichž jeden odpovídá chování MLA proteinu.⁹⁷ K analýze byl dále vybrán dříve diskutovaný pruh odpovídající velikostí přibližně dimeru MLA, který se z SEC purifikovaných variant MLA objevoval pouze u MLA092, kde měl nezanedbatelnou koncentraci. Ačkoliv nebyly „dimerní“ pruhy pozorovány u žádné z jiných SEC purifikovaných variant, vyskytovaly se poměrně častěji u Ni-NTA purifikovaných proteinů určených pro původní stanovení vazebné afinity, jak je vidět v **tabulce 7**.

Konkrétní analyzované pruhy SDS-PAGE gelu jsou zobrazeny na **obrázku 17**.



Obrázek 17 – Barvený SDS-PAGE gel SEC purifikovaných zakonzentrovaných frakcí His-MLA092-V5. Červeně jsou vyznačeny pruhy analyzované metodou MS.

Výsledky MALDI-TOF analýzy prokázaly jednoznačně identitu pruhu 1 (SEMS frakce) jako bakteriální Hsp70 protein (pokrytí 76,3 %). Toto vysvětlení koresponduje se všemi pozorovanými vlastnostmi. Na gelové permeační chromatografii tento protein putuje jako komplex, na SDS-PAGE se rozpadá na samotné Hsp70 a na MLA protein (jak ukazují SDS-PAGE a ELISA experimenty).

Pruh 2 má pouze 37,4 % pokrytí sekvence His-MLA092-V5, sekvenci ale odpovídal velmi intenzivní pík s m/z 1627,7467. Tento pruh tedy odpovídá dimernímu myomedinu, jak naznačuje i změřená molekulová hmotnost odpovídající dvojnásobku hmotnosti monomeru. Podobné výsledky byly pozorovány i u MALDI-TOF analýzy obdobného pruhu proteinu His-TEV-MLB049-V5 z jiné myomedinové knihovny (výsledky nepublikovány).

Na základě všech zmíněných skutečností bylo zjištěno, že SEMS frakce proteinů sestávají v gelové permeační chromatografii z komplexu Hsp70-MLA, na SDS-PAGE se rozpadají na konstituentní proteiny. Takto vzniklý MLA protein prokazoval v ELISA experimentech MLA024 a MLA025 příznivé vazebné vlastnosti, jednalo se tedy minimálně v těchto případech o funkční vazebný protein.

Existence dimerů je na základě výsledků velmi pravděpodobná, jejich charakter je však neznámý. V zde uvedené sérii experimentů byl dimer pozorován pouze pro jednu variantu, pro kterou v předchozí sérii experimentů (byť purifikovaných jen pomocí Ni-NTA sefarysy) nebyl takto výrazný pruh na SDS-PAGE pozorován (**tabulka 7, MLA092**). Obdobné chování, tedy výskyt dimeru jen u některých variant, bylo pozorováno i u MLB knihovny (odlišná Myomedinová knihovna cílicí na PGT 126 bnAb). Dimer putuje v denaturujících podmínkách SDS-PAGE jako intaktní protein, v podmínkách gelové permeační chromatografie se projevuje nesymetričností píku SELS (monomerní myomedin). Charakter vzájemné vazby dvou monomerů musí být tedy buď kovalentní, nebo charakteru, který není narušen v podmínkách SDS-PAGE elektroforézy. Zároveň se zdá, že je vznik dimeru závislý na podmínkách.

5.1.9.4 DLS měření

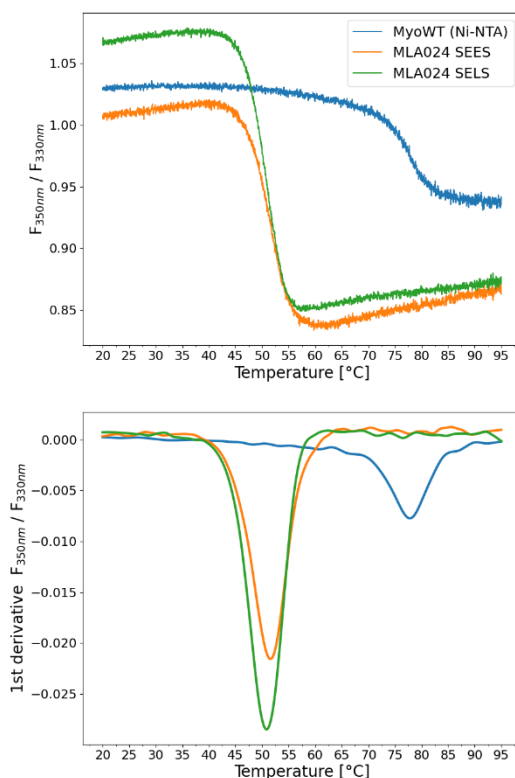
Bylo provedeno orientační měření DLS na frakcích, kterých se podařilo získat dostatečné množství pro analýzu. Hydrodynamický průměr monomerních forem myomedinů (frakce SELS), získaný přibližně odečtem maxima píků distribuce procentuálního objemu zabraného jednotlivými velikostmi částic, byl stanoven na $D_H=3,357-3,903$ nm pro MLA024 (měřeno celkem ve 3 experimentech s pokaždé znovu produkovaným a purifikovaným proteinem), MLA025 (měřeno ve dvou experimentech), MLA092 a MyoWT proteiny.

5.1.9.5 DSF měření termostability

Pomocí diferenční skenovací fluorimetrie na přístroji Prometheus NT.48 byla měřena teplotní stabilita některých purifikovaných frakcí. Výsledné naměřené hodnoty T_m získané z inflexních bodů poměrů fluorescence tryptofanů při 350 a 330 nm) jsou uvedeny v **tabulce 11**. Pro frakci SELS MLA092 chybí původní data a křivky průběhů zde tak nejsou zobrazeny.

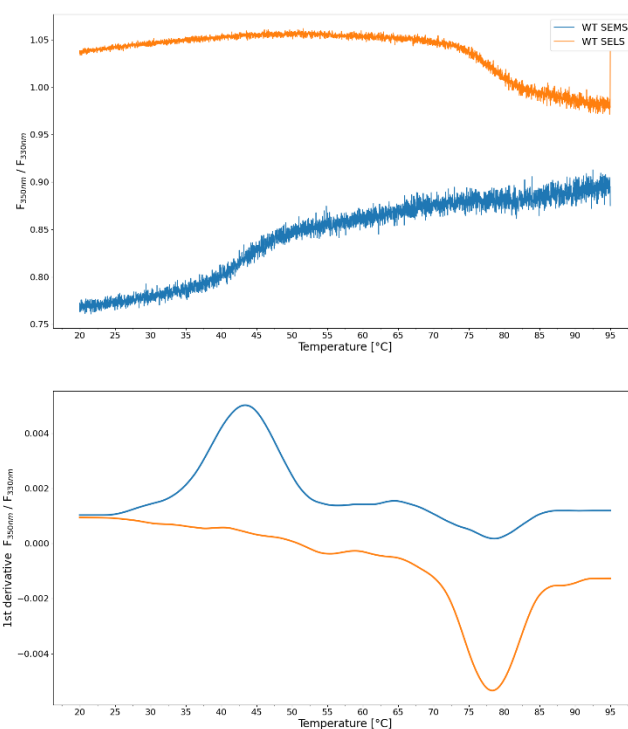
Tabulka 11 - T_m hodnoty získané měřením teplotní stability proteinových frakcí metodou DSF.

His-MLA-V5 frakce	$T_m / ^\circ\text{C}$
MLA024 SEES	51,5
MLA024 SELS	50,8
MLA092 SELS	41,6
MyoWT (Ni-NTA)	77,7
MyoWT SEMS	43,3
MyoWT SELS	78,3
MLA158 SEMS	45,3
MLA158 SELS	51,3



Obrázek 18 –Průběh poměrů fluorescencí tryptofanů při 350/330 nm získané měřením DSF během zahřívání roztoku frakcí proteinu MLA024 SEES a SELS a Ni-NTA purifikovaného MyoWT. Nahoře) Poměry intenzit emisí při 350 nm a 330 nm v závislosti na teplotě Dole) První derivace poměrů intenzit emisí v závislosti na teplotě

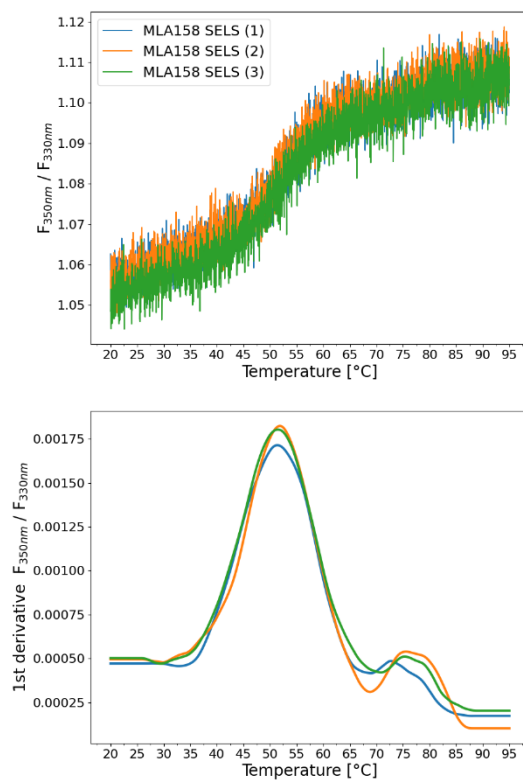
Obrázek 18 zobrazuje výsledné křivky DSF měření pomocí Ni-NTA purifikovaného MyoWT a dvou frakcí MLA024 (oligomerní a monomerní). Je vidět, že je teplotní stabilita obou MLA024 frakcí snížena oproti MyoWT (inflexní bod snižen ze 77,7 °C na 50,8-51,5 °C). Zároveň je patrné, že je bod tání T_m pro obě frakce MLA024 velmi blízký, to dobře koresponduje s ostatními experimenty a podává další důkaz o oligomerním charakteru frakce SEES (obdobné chování jako monomer).



Obrázek 19 – Průběh poměrů fluorescencí tryptofanů při 350/330 nm získané měřením DSF během zahřívání roztoku frakcí proteinu MyoWT SEES a SELS. Nahoře) Poměry intenzit emisí při 350 nm a 330 nm v závislosti na teplotě Dole) První derivace poměrů intenzit emisí v závislosti na teplotě. Oranžová křivka popisuje frakci SELS, modrá frakci SEES.

Interpretace výsledků měření frakcí SELS a SEES MyoWT, vizualizovaná na **obrázku 19**, je následující. Hodnota $T_m = 78,3 \text{ }^\circ\text{C}$ stanovená v tomto experimentu odpovídá dříve stanovené hodnotě $77,7 \text{ }^\circ\text{C}$, měřené na MyoWT purifikovaném pouze metodou s využitím Ni-NTA agarosy. V získaných MyoWT roztocích je monomer naprosto majoritně zastoupen, viz **obrázek 15**. Křivka odpovídající frakci SEES obsahuje dva přechody. Druhý z nich má identické $T_m = 78,3 \text{ }^\circ\text{C}$ a odpovídá tedy monomernímu myomedinu z frakce SELS. První přechod se nachází při teplotě zhruba $45 \text{ }^\circ\text{C}$ a odpovídá bakteriálnímu Hsp70. Tento experiment je tedy v dobrém souhlasu s dříve diskutovaným popisem charakteru frakce SEES (Hsp70+Myomedin).

Výsledky měření termostability frakce SELS proteinu MLA158 jsou uvedeny v **obrázku 20**. Je vidět výrazný přechod při $51,3 \text{ }^\circ\text{C}$ a minoritní přechod mezi $75\text{--}85 \text{ }^\circ\text{C}$. Svojí pozicí odpovídá tento přechod nápadně přechodu při zahřívání MyoWT. Je možné, že se jedná o konformační změnu, která je společná oběma těmito proteinům při této teplotě.



Obrázek 20 – Průběh poměrů fluorescencí tryptofanů při 350/330 nm získané měřením DSF během zahřívání roztoku SELS frakce proteinu MLA158. Nahoře) Poměry intenzit emisí při 350 nm a 330 nm v závislosti na teplotě Dole) První derivace poměrů intenzit emisí v závislosti na teplotě. Měření proběhlo ve třech opakováních.

5.2 Charakterizace proteinových variant metodami *in silico*

5.2.1 Celkový cíl

Výpočetní část této práce byla motivována snahou vysvětlit určité projevy selektovaných proteinových variant, které značně stěžovaly jejich charakterizaci biofyzikálními metodami, především značná heterogenita produkovaného proteinu, rozdílné chování jednotlivých purifikovaných frakcí včetně frakce vystupující jako oligomer. Úsilí oddělit jednotlivé frakce a ty dále charakterizovat komplikovala rovněž různá stabilita a tendence nabývat různých forem či konformaci především při koncentrování proteinu tak, aby bylo vyhověno požadavkům dané instrumentální metody.

Dosud neexistují experimentálně vyřešené struktury komplexů MLA proteinů s 10E8 protilátkou, díky problémům s produkcí, čistotou a nedostatkem protilátky. Stejně obtíže znemožňují jedna provést standardní biologicko-strukturní experimenty jako je např. rentgenová krystalografie, tak stanovení vazebné afinity proteinů biofyzikálními metodami jako je kalorimetrie či termoforesa. Zároveň existuje jen omezený vhled do samotné mimikující aktivity proteinů a do role jednotlivých reziduí. *In silico* metody byly vybrány jednak pro částečné obejítí zmíněných experimentálních obtíží tak pro získání atomistického popisu mimikující aktivity. Znalost mimikujícího mechanismu by mohl vést vylepšování vazebných proteinů.

Konkrétním cílem této práce bylo získání atomistického pohledu na vazbu MLA proteinů s 10E8, získat fyzikálně-chemické znalosti o jejich interakci a vysvětlit mechanismus mimikování.

Obecnějším cílem výpočetní části této práce bylo vytvoření protokolu, sloužícího k *in silico* analýze vazebných interakcí proteinů problematických z experimentálního hlediska, bez vyřešených struktur a přesných experimentálních stanovení.

5.2.2 Homologní modelování

Pro účely homologního modelování bylo využito GUI rozhraní Modeller 9.2.4. v UCSF Chimera. Použity byly celé sekvence myomedinových proteinů včetně tagů, ke kterým byly pomocí algoritmu BLAST nalezeny homologní struktury. Oproti běžnému postupu využití více templátů při homologním modelování byl zvolen jen jeden templát – PDB:3RBS, řetězec A. Tento byl zvolen z důvodu nejvyšší míry homologie a kvality struktury. Použití více templátů zároveň nevedlo ke zlepšení výsledných prvotních modelů.

Ke každé variantě Myomedinů bylo vygenerováno 5 modelů. Podle vizuální inspekce byly mezi modely podstatné rozdíly jen v orientaci terminálních tagů, které byly později odstraněny. Proteinový hlavní řetězec měl mezi modely jen nepatrné změny, soudě dle zDOPE skóre kvality (více negativní je lepší, **tabulka 12**). Homologní model MyoWT byl vytvořen primárně jako kontrola kvality celého procesu, alternativou by bylo jeho získání skrz úpravu templátové struktury. Tento první krok byl prováděn za přítomnosti terminálních tagů, které se později ukázaly pro celý proces jako problematické a byly tedy odstraněny „ořezem“ v aplikaci PyMol.

Tabulka 12 – Skóre kvality, sekvenční homologie a vybraných modelů z pěti homologních modelů vytvořených na základě templátu 3RBS (řetězec A) Modellerem. Skóre jsou vypočtena za přítomnosti terminálních tagů.

Var.	GA341	zDOPE	Sekvenční id. %	Vybraný model
MLA024	1.00	-0.45	91.9	1
MLA025	1.00	-0.55	90.1	1
MLA092	1.00	-0.49	89.2	1
MLA158	1.00	-0.16	90.1	1
MyoWT	1.00	-0.91	100	3

V UCSF Chimera byly zároveň vygenerovány Ramachandranovy výnosy pro kontrolu kvality prvotních modelů z hlediska dihedrálních úhlů peptidové vazby. Stejný výnos byl vytvořen i pro templátovou strukturu 3RBS. Výsledné výnosy jsou zobrazeny v **příloze B 3**. Bylo patrné, že většina dihedrálních úhlů leží v oblastech ideálních a dovolených

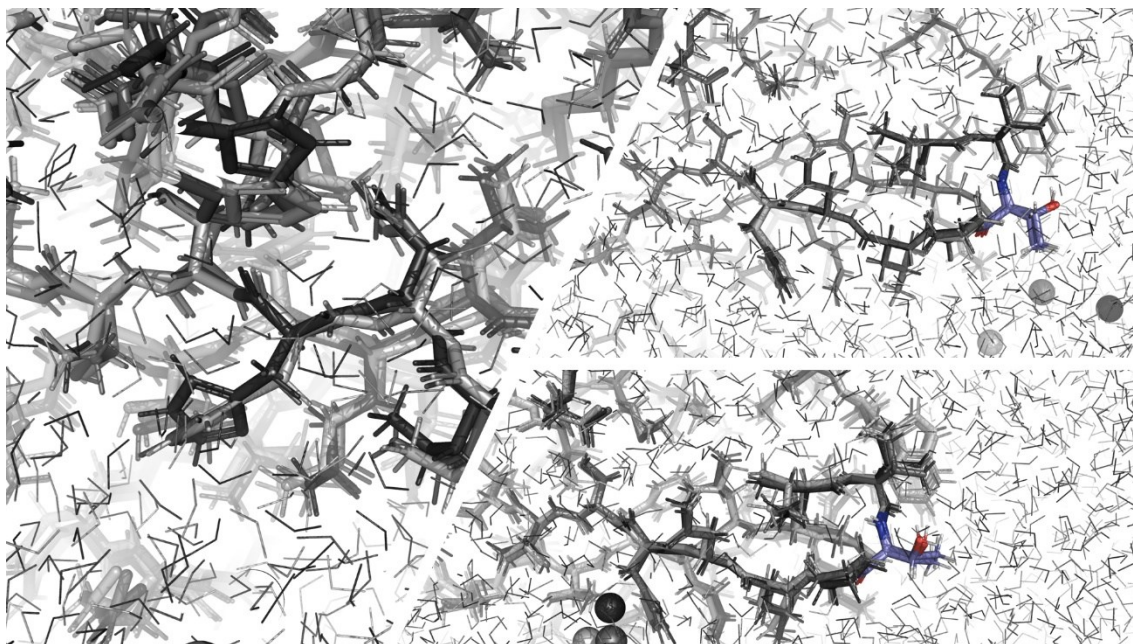
kombinacích, jen několik málo aminokyselin mělo své peptidické dihedrální úhly mimo tyto oblasti.

Pro účely MD simulací byly pomocí skriptů ProPKa vypočteny teoretické hodnoty pKa jednotlivých reziduí (výsledky nejsou součástí této práce). Pokud by výsledky naznačovaly takový rozdíl mezi tabelovanou a vypočtenou hodnotou, že by se změnil protonační stav rezidua při pH 7, byly by pro přípravu struktur do simulací tyto protonační stavy manuálně zavedeny do struktury (pro simulaci fyziologických podmínek). Tento případ nenastal pro žádný z modelů.

Zajímavým výsledkem byly teoretické hodnoty optim pH stability proteinů. V současné experimentální praxi je stabilita proteinů sub-optimální a tyto výsledky by mohly sloužit jako odrazový můstek pro optimalizaci pufrů a expresních podmínek. Vypočtené hodnoty pH optim byly následující – MLA024 8,4 ; MLA025 8,3 ; MLA092 9,4 ; MLA158 8,7 ; MyoWT 8,5.

5.2.3 Vylepšení a validace modelů (MD simulace)

Dalším krokem bylo vylepšení homologních modelů z hlediska geometrie a energetiky. Byl zvolen unikátní protokol popsáný v **Kapitole 4.2.3.1** založený na prvním kroku kombinované validace/vylepšení pomocí standardní molekulové dynamiky, následované kontrolou modelu a dalším krokem molekulové dynamiky s restrainovanými hlavními řetězci. Posledním krokem aplikovaným jen na některé modely bylo cílené zlepšení konkrétních reziduí pomocí nerovnovážné simulace s mražením. Vizuální přehled jednotlivých kroků zobrazuje **obrázek 21**. Celková doba simulací aplikovaná na modely byla mezi 30–35 ns (podle toho, zda bylo provedeno poslední 5 ns kolo) – hodnota snadno dosažitelná na moderním hardwaru. Celkový proces je zároveň do velké míry automatizovatelný a jak bude diskutováno dále, přinesl ve zkoumaném případě výrazná zlepšení homologních modelů. Protokol navíc přímo navazuje na standardní postup – MD simulace pro validaci homologních modelů. I tyto krátké simulační časy stačily pro dosažení konvergenčí RMSD a gyračních poloměrů.



Obrázek 21 – Vylepšování homologních modelů pomocí krátkých MD simulací. Barevně je zvýrazněno residuum zájmu. Tmavší odstíny zobrazují subsekventní kroky MD simulace.

- Vlevo) První krok Standardní MD simulace
 Vpravo dole) Druhý krok MD simulace s restrainovaným hlavním proteinovým řetězcem
 Vpravo nahoře) Třetí krok MD simulace s mražením cílená na specifická residua

Cílem první fáze MD simulací bylo provést energetickou minimalizaci, validaci a získání souboru konformací vycházejících z prvotních modelů. Vzhledem k zamýšlené povaze celého protokolu a technickým omezením byly časy produkčních simulací pouhých 20 ns. Z hlediska dosažení konvergence klíčových vlastností (RMSD, Rg) se však jednalo pravděpodobně o čas dostačující. Samotné struktury navíc v rámci další práce prošly dalšími koly MD simulací.

Na základě výsledků těchto geometrických analýz a dalších cílů práce byly po prvním kole MD simulací prvotních homologních modelů ze struktur odstraněny terminální tagy. Získat nativní sbalenou strukturu tagů, pokud vůbec existuje, by bylo v používaném uspořádání prakticky nemožné vzhledem ke krátkým časovým škálám. Tagy by zároveň mohly interferovat s dokováním. Z hlediska propojení teoretických výpočtů a experimentálních by nebylo vhodné stavět závěry na základě pouze jednoho typu konstrukcí (zde s N-terminálním 6×-His-Tagem a C-terminálním V5-tagem), jelikož jak v *in vivo* tak v *in vitro* experimentech byly používány i konstrukty s odlišnými tagy, bez tagů či fúzované s jinými proteiny. Cílem bylo zároveň zjistit chování samotného

myomedinu, nikoliv artificiálního konstruktu. Toto odstranění tagů přineslo určitou neintuitivitu číslování struktur, která bude zřejmá v pozdějším textu. Odstraněním N-terminálního 13 aminokyselin dlouhého konce bez přečíslování vznikly struktury začínající Lys14. Toto číslování bylo nadále zachováno a bude používáno, pokud nebude uvedeno jinak.

Z posledních 10 ns trajektorií. kvůli dosažení konvergenčí, byly vybrány reprezentativní struktury/centroidy.

Kontrola výstupu tohoto a dalších vylepšovacích kroků proběhla pomocí balíku skriptů MolProbity, které sledují klíčové vlastnosti proteinových modelů najednou. Tento balík zároveň optimalizuje pozice vodíků a 180° rotace („flipping“) některých reziduí pro zlepšení energetických vlastností modelu. Funkce optimalizace vodíků byla využita pro všechny dále zmíněné modely, funkce rotace reziduí pouze pro dva modely z prvního kole vylepšování – MLA024 (Asn43), MLA025 (Gln109), MLA092 (Gln86).

Pro všechny modely bylo provedeno i druhé kolo vylepšování se vstupními strukturami z minulého kola, zaměřené primárně na postranní řetězce, ve kterých byly dle MolProbity, koncentrovány problémy s délkami a úhly vazeb. Hlavní řetězec byl restrainován silou $1000 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-1}$ a byla simulována 10 ns dlouhá MD trajektorie. Z této trajektorie byly obdobným způsobem vybrány reprezentativní konformace, tentokrát však, na základě porovnání centroidů/průměrů dle MolProbity skóre, byly vybrány průměrné konformace. Vzhledem k tomu, že nejsou průměrné struktury „fyzikální“ a nejsou skutečnou součástí MD trajektorie – nemusí zobrazovat reálnou fyzikální konformaci molekuly. Proto byly tyto struktury před další analýzou nejprve energeticky minimalizovány kroky identickými s ostatními popsányi MD simulacemi.

Výsledky některých klíčových parametrů analýzy kvality modelů z prvního a druhého kola vylepšování jsou sumarizovány v **tabulce 13**.

Tabulka 13 – Výsledky geometrické analýzy homologních modelů mezi prvním a druhým kolem vylepšování pomocí MD simulací. K získání analýz bylo použito balíku skriptů MolProbity. Pro každý parametr levý sloupec zobrazuje hodnoty z prvního kola, pravý sloupec z druhého kola.

Varianta	Překryvy atomů		Příznivé rotamery %		Příznivé dih. úhly %		Špatné vazby / úhly		MolProbity skóre / percentil	
MLA024	0	0	88,17	97,85	94,50	99,08	19/48	1/1	0,90 / 100	0,50 / 100
MLA025	0	0	93,26	97,75	91,74	93,58	16/53	0/3	1,26 / 99	0,96 / 100
MLA092	1,11	0	90,00	92,22	94,50	93,58	12/59	2/4	1,59 / 93	1,32 / 98
MLA158	0,54	0	89,01	93,41	96,33	97,25	29/58	1/0	0,96 / 100	0,64 / 100
MyoWT	1,12	0	92,31	93,41	94,50	99,08	31/57	3/1	1,46 / 96	0,53 / 100

Jak je patrné, druhé kolo, dle používané metriky, přispělo k vylepšení homologních modelů výraznou měrou. Drobná zlepšení zaznamenaly u všech modelů jak počty „správných“ rotamerů tak počty aminokyselin s dihedrálními peptidickými úhly v ideálním rozmezí na Ramachandranově výnosu (s výjimkou varianty MLA092, kde naopak došlo k drobnému zhoršení). Primárním cílem druhého kola vylepšování bylo odstranění špatných délek a úhlů vazeb postranních řetězců. Toho bylo dosaženo. u všech variant byly počty těchto nedokonalostí sníženy na nízké jednotky. Celkové vylepšení je vidět i na kumulativních MolProbity skóre (menší znamená lepší), kde došlo u všech variant ke zlepšení.

Dalším aplikovaným kritériem kvality byla správnost záměnných aminokyselin, které by měly být důležité pro vazbu s 10E8. Pohled na konkrétní aminokyseliny z hlediska špatných vazebných délek a úhlů je uveden v **tabulce 14**. Na základě těchto výsledků byly vybrány varianty MLA025 a MLA092 pro další vylepšení, směřované specificky na oblasti uvedených residuí.

Tabulka 14 – Špatné vazebné úhly a délky modelů po druhém kole vylepšení. Zvýrazněny tučně jsou pozice odpovídající randomizovaným postranním řetězcům.

Varianta	Špatné vazebné úhly	Špatné vazebné délky
MLA024	His60	Lys106
MLA025	Thr35 (CA-C-N) , Phe46, Asn47	X
MLA092	Glu48, Glu78, Thr91 (CA-C-N) , His95	Met70
MLA158	X	Lys106
MyoWT	Lys24, Lys106	Lys24, Lys74

Bylo provedeno třetí kolo vylepšení pro modely MLA025 a MLA092. Směřované bylo na aminokyseliny uvedené v **tabulce 14**. Pro získání souboru možných konformací postranních řetězců byly provedeny nerovnovážné MD simulace. Oproti druhému kolu nebyl hlavní řetězec proteinu restrainován ale zmražen – bylo kompletně zabráněno jeho pohybu v rámci simulace. Zmražení bylo aplikováno i na všechna kromě uvedených residuí tak, že zůstala nezmražena vždy uvedená aminokyselina a z každé strany jedna sousední aminokyselina. Všechny vazby s vodíkovými atomy byly constrainovány algoritmem SETTLE (standardní postup). Speciální pozornost byla věnována správné tvorbě indexů částic pro zmražení, které musely zahrnovat správnou volbu částic, tedy konkrétních aminokyselin. Zároveň nesměly obsahovat vodíky, vazby na ně, molekuly solventu a iontů. V rámci ekvilibrace byly zároveň nezmražené aminokyseliny restrainovány.

Bylo provedeno po 5 ns MD simulací. Z těchto simulací byly vybrány reprezentativní centroidy s nízkou hodnotou cut-off (jen minimální změny geometrií během MD simulací), které byly hromadně analyzovány pomocí MolProbity. Vybrán byl vždy nejlepší centroid, reprezentující konformaci s co nejvíce vhodnými vazebnými vzdálenostmi a úhly. Konečné výsledky třetího kola zlepšování modelů ve srovnání s druhým kolem jsou uvedeny v **tabulce 15**.

Tabulka 15 - Výsledky geometrické analýzy homologních modelů mezi druhým a třetím kolem vylepšování pomocí MD simulací. K získání analýz bylo použito balíku skriptů MolProbity. Pro každý parametr levý sloupec zobrazuje hodnoty z druhého kola, pravý sloupec ze třetího kola.

Varianta	Překryvy atomů		Příznivé rotamery %		Příznivé dih. úhly %		Špatné vazby / úhly		MolProbity skóre / percentil	
MLA025	0	0	97,75	97,75	93,58	93,58	0/3	0/2	0,96 / 100	0,96/100
MLA092	0	0	92,22	93,33	93,58	93,58	2/4	0/5	1,32/98	1,38/97

Dle výsledků třetí kolo analýzy nepřineslo výrazné zlepšení, které by se odrazilo na „makroskopických“ vlastnostech struktury jako celku, některé vlastnosti se dokonce mírně zhoršily. Klíčovým zlepšením však byla eliminace nevhodných vazebných délek a úhlů v oblastech aminokyselinových záměn.

Celý výše popsaný proces, od prvotních homologních modelů po třetí kolo vylepšování MD simulací, sloužil dvěma hlavními účelům – jako „proof-of-concept“ vylepšování homologních modelů krátkými MD simulacemi a z důvodu následujícího kroku rigidního dokování. Speciální ohled byl brán na správnost modelu aminokyselinových záměn.

S ohledem na první cíl („proof-of-concept“) je dle použitých kritérií evidentní výrazné zlepšení modelů po dvou kolech simulací a odstranění specifických chyb u dvou variant po třetím kole. Z hlediska celkových skóre kvality byly všechny modely v rozmezí percentilů 97-100. Popsaný protokol přinesl výrazně lepší výsledky než dva vyzkoušené standardní nástroje pro vylepšování homologních modelů (viz **kapitola 4.2.3.1**). Celý proces může být jednoduše implementován i na jiné systémy.

S ohledem na druhý cíl (příprava modelů pro rigidní dokování) bylo dosaženo vytvoření kvalitních homologních modelů.

5.2.4 Dokování a prvotní výběr modelů

K dokování typu protein-protein byl vybrán webový server ClusPro2 provádějící *ab initio* dokování, pro porovnání a kontrolu byl používán webový server HDock. Jako omezená kontrola celého procesu od dokování po další kroky (viz dále) sloužilo porovnání vyřešené struktury komplexu 10E8Fab-MPER (PDB:4G6F) a dokováním získaných struktur (dokování samotného 10E8Fr. s MPER). ClusPro2 je rozšířený, standardní nástroj pro dokování protein-protein. Jeho výhoda spočívá zejména v samplování velkého počtu konformací, následného výběru reprezentativních konformací z clusterů a jejich automatické energetické minimalizaci. Toto je prováděno interně pomocí balíku PIPER.⁹⁸ Další výhodou jsou metody skórování vyvinuté pro konkrétní typy komplexů (zde bylo použito skórování pro komplexy s protilátkou), možnosti „stínění“ reziduí od vazby a dokování na základě experimentálních SAXS křivek. Bylo prokázáno, že 30 největších clusterů získaných ClusPro dokováním obsahuje konformace blízké nativním pro 92 % komplexů z testovací sady protein-protein struktur, úspěšnost může být dle autorů vyšší pro některé typy komplexů. Standardní postup zahrnuje výběr deseti nejvíce populovaných clusterů.⁷⁷ Tento postup byl zvolen i v této práci.

Vstupními strukturami dokování byly v předchozích krocích získané vylepšené homologní modely a zkrácené Fab fragmenty 10E8 protilátek (10E8fr.). Zkrácení sloužilo ke snížení velikosti struktury (pro zrychlení následných MD simulací) tak, že z nativní struktury protilátky zůstaly pouze variabilní VH a VL úseky a krátké nestrukturované sekvence pojící tyto úseky v nativní struktuře ke konstantním CH1 a CL úsekům.

V této práci nebyl, použit žádný „bias“, jakým by bylo například omezení vazby jen na oblast aminokyselinových záměn. Vazba byla zkoumána zcela *ab initio* bez stínění reziduí, předem známých vzdáleností či SAXS křivek. Tento přístup byl odlišný od přístupu popsáném v článku naší laboratoře o MLA vazebných proteinech.⁵¹

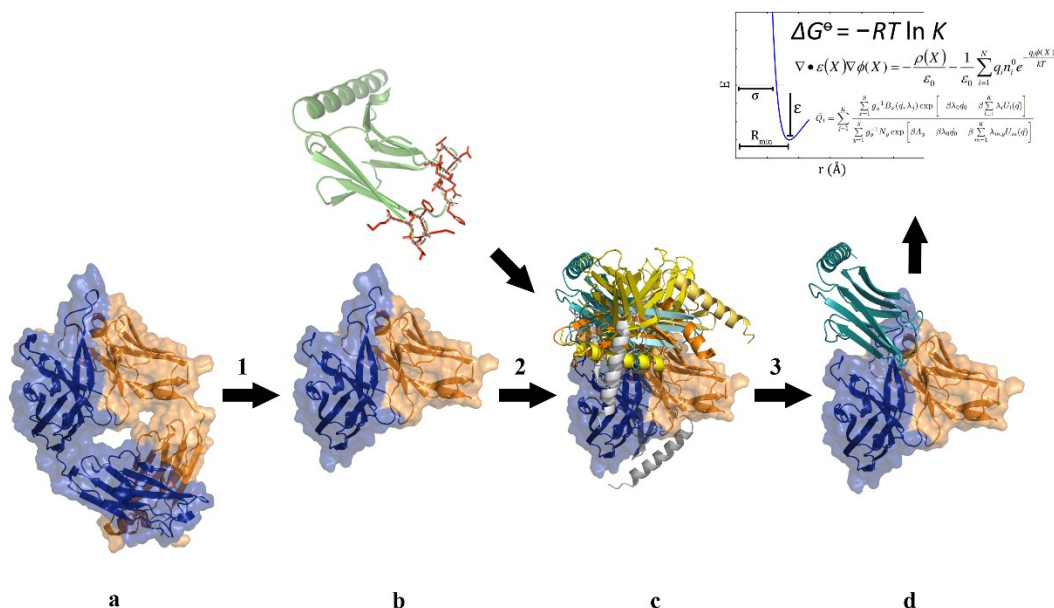
Vybraná metoda dokování používá aproximaci rigidních těles. Jedná se o standardní postup, který však může v některých případech minout důležité konformační změny během procesu navazování. Určité vylepšení přináší MD simulace, které přináší dynamičnost do tohoto „statického pohledu“ na interakce. Vzhledem k relativně krátkým

simulačním časům používaným v této práci není možné toto zlepšení uvažovat, celá práce tedy předpokládá vazbu dle modelu „zámek a klíč“ – uvažuje jen minimální konformační změny při navázání proteinů na cílové fragmenty protilátky. Dalším vstupním předpokladem dokování bylo zanedbání vázaných molekul vody pro vazbu. Tento předpoklad byl dán odstraněním všech molekul vody ze struktury protilátkového fragmentu (tato část postupu byla identická s publikovaným článkem).

Dalším faktorem hodným zřetele je fakt, že i přes poměrně účinné moderní dokovací algoritmy, nemusí identifikovaná struktura s nejnižší hodnotou volné Gibbsovy energie (přibližně aproximované „dokovacím skóre“) odpovídat struktuře nativní. Pro plnou validaci výsledků dokování by bylo nutné provést experimentální stanovení. Z metod, potenciálně aplikovatelných do budoucna (které nejsou součástí této práce) se nabízí například porovnání teoretické a experimentálně získané křivky SAXS, případně experimentální získání některých vzdáleností komplexu metodou FRET a jejich následné využití jako vodítka pro dokování a molekulovou dynamiku. Vzhledem k nedostupným materiálům jsou tyto experimentální metody prozatím vyloučeny.

Pro tuto práci bylo ze struktur získaných dokováním ClusPro2 zvoleno 10 nejpopulovanějších clusterů, vybraných dle protilátkového skórovacího módu. Modely byly vizuálně zkontrolovány – ořezem 10E8 Fab fragmentu (viz **kapitola 4.2.3.2**) vznikl ne-nativní povrch, který není ve skutečné molekule přítomen. Kontrola sloužila ke zjištění, jestli některé modely nezahrnují vazbu s tímto povrchem. Na základě této kontroly byly vyřazeny modely 9 pro MLA024; 7 pro MLA025; 7 pro MLA092.

Schematicky zobrazený postup dokování a následných kroků ukazuje **obrázek 22**.



Obrázek 1 – Schéma tvorby a výběru vazebného modelu myomedinových proteinů s protilátkovým fragmentem 10E8.

a) Fab fragment 10E8 (PDB:4G6F)

1) Ořez Fab fragmentu na protilátkový fragment sestávající z variabilních oblastí protilátky

b) Vstupní struktura protilátky pro dokování

2) Rigidní dokování s homologními modely proteinů MLA

c) Výběr deseti nejpopulovanějších clusterů. Ve spodní části je vidět interakce, ke které by v nativní protilátce nemohlo dojít – takové modely nebyly uvažovány.

3) Výběr jediného modelu a jeho analýza – kroky MM/PBSA a Umbrella sampling, analýza intermolekulových interakcí

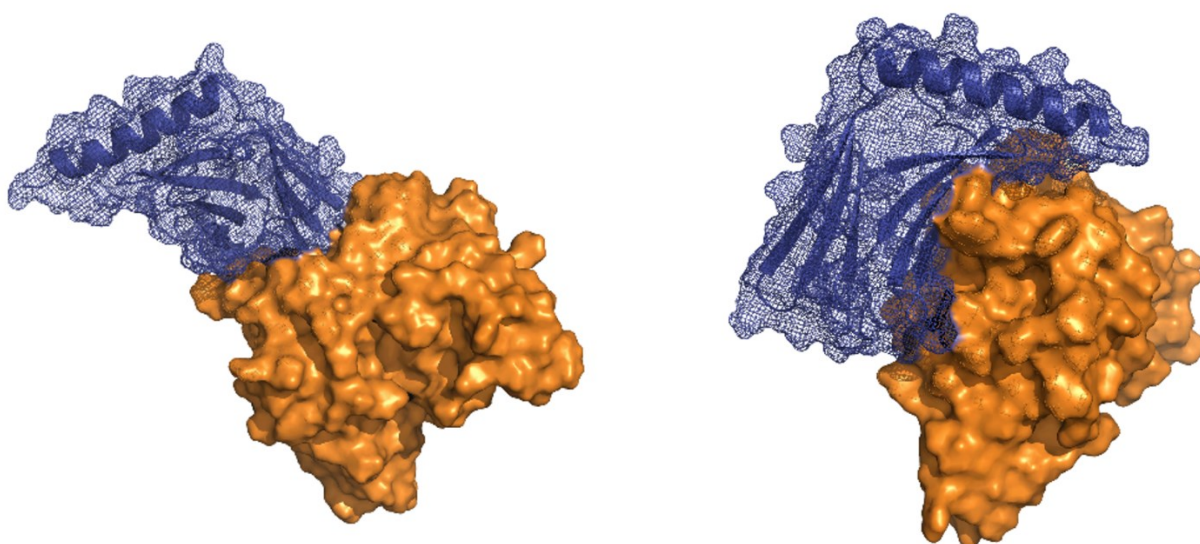
4) Jediný vybraný model pro jednu variantu MLA

5.2.4.1 β -vazebný mód

Vizuální inspekce modelů získaných dokováním přinesla překvapivý závěr – žádný ze získaných modelů, s výjimkou dvou modelů varianty MLA158 (modely 2 a 3) neodpovídal doposud předpokládanému vazebnému módu, ve kterém by mělo k vazbě docházet převážně skrze zaměněné aminokyseliny. Jejich důležitost je prokázána *in vivo* experimenty, ve kterých byly patrné rozdílné neutralizační schopnosti proteinů na základě těchto aminokyselin. Toto pozorování znamená, že musí nějakým způsobem opravdu docházet ke kontaktu zaměněných aminokyselin s protilátkou. Určitou nejistotou *in vivo* studií je, zda-li nedochází k neutralizaci pomocí jen jedné, specifické vazebné konformace (která může existovat jen tranzientně) či zda-li vůbec dochází k neutralizaci pomocí indukce tvorby protilátek 10E8 (používaných v této práci) a ne jiných, podobných

protilátek (např. 4E10). Jak bude vidět v dalším textu, β -vazebné módy mohou být termodynamicky stabilní a mít vysoce příznivé vazebné energie, včetně interakcí aminokyselinových záměn s klíčovými rezidui 10E8. Dosud předpokládaný „**smýčkový vazebný mód**“ byl analyzován u modelů 2 a 3 proteinu MLA158. Pro kontrolu, zda není tento nečekaný mód pouze artefaktem provedeného dokování bylo dokování se stejnými vstupními strukturami provedeno i pomocí serveru HDock, kde byly podobně ve velké převaze právě tyto vazebné módy (výsledky nezveřejněny).

Další kapitoly budou detailněji popisovat zejména modely obsahující tento β -vazebný mód. Prozatimní základní obraz o odlišnostech popisuje **obrázek 23**. Oproti dosud předpokládanému smýčkovému vazebnému módu řízeného zejména aminokyselinovými záměnami na nestrukturovaných smýčkách má dokováním preferovaný β -vazebný mód viditelně větší interakční tvořený ze strany myomedinu rezidui β -listů.



Obrázek 23 – Dva vazebné módy myomedinových proteinů na protilátkový fragment 10E8. Oranžově je zobrazen rozpouštědlu dostupný povrch (SASA) protilátkového fragmentu 10E8. Modře je zobrazena síť SASA povrchu myomedinu a jeho sekundární struktura.

(Vlevo) Smyčkový vazebný mód. Skrze menší interakční povrch je kontaktována protilátka záměnami.

(Vpravo) β -vazebný mód. Větší interakční povrch, řízený rezidui β -listů proteinu.

V prvním přiblížení by zvětšený interakční povrch β -vazebného módu mohl znamenat zvýšenou sílu interakce. Stejně struktury se s výrazným přesahem objevovaly i v práci

Ing. Jiřího Černého, PhD. z *Laboratoře strukturní bioinformatiky proteinů Biotechnologického ústavu AV ČR*, který využíval stejné dokovací metody (ClusPro2) avšak vycházel z odlišných homologních modelů. To může částečně vylučovat vliv homologních modelů použitých v této práci (nezveřejněné výsledky).⁹⁹

Specifika konformace β -vazebného módu budou diskutována v **kapitole 5.2.9**, již nyní je však nutno zmínit, že se tento vazebný mód nesestává z jediné konformace, ale obsahuje několik pozorovaných konformací s odlišnou pozicí proteinů vůči sobě, které však mají společnou interakci přes rezidua β -listů myomedinů.

5.2.5 MM/PBSA a MM/GBSA analýza

Pro „hrubý výběr“ mezi deseti dokováními získanými modely komplexů myomedinových proteinů s 10E8fr. bylo využito metody MM/PBSA. Pro získání vstupních trajektorií byly pro každý model provedeny 20 ns MD simulace se silovým polem CHARMM36m identickým postupem simulacím z prvního kola vylepšování homologních modelů. Ze všech simulací bylo vybráno posledních 5 ns na základě dosažených konvergencí parametrů RMSD a gyračních poloměrů. Grafy průběhu těchto parametrů v simulacích jsou zobrazeny v **příloze B 4 a B 5**.

Tato metoda leží na střední pozici v přesnosti mezi nejméně přesnými metodami skórování a nejpřesnějšími metodami „alchemical perturbations“. Metoda je i přes zvýšenou přesnost oproti dokovacím skóre zatížena řadou aproximací (zanedbání entropie, řešení solvatačních energií, využití empirických silových polí, implicitní solvatace ad.), je však velmi vhodná pro porovnání, nikoliv kvantitativní popis, vazebných energií ligandů na stejného vazebného partnera, k čemuž byla využita i zde.^{100,101}

Využito bylo nejběžnější implementace přístupu MM/PBSA s využitím jedné simulace komplexu (oproti teoreticky méně přesnému přístupu tří simulací komplex, ligand, protein). K výpočtům bylo využito posledních 5 ns dříve popsaných MD trajektorií (celkem 50 snímků). K popisu systémů bylo využito konverze struktur z CHARMM36m do ff99-SB silového pole. K výpočtu solvatačních energií bylo využito Poisson-

Boltzmannovy rovnice (polární část, parametr $\epsilon_{\text{int}} = 2$ zvolen dle obvyklé hodnoty a „průměrně nabitého interakčního povrchu“, otestovány byly i hodnoty 1 a 4) a metody SASA (nepolární část, korelace mezi solventu dostupných povrchem a nepolární solvatační energií).

Samotné výpočty proběhly pomocí balíku `gmx_MMPBSA` a validací pomocí balíku `g_mmpbsa` (výsledky získané ním nejsou zveřejněny, byly však v relativní shodě).

Více korektním postupem provedení MM/PBSA analýz by bylo „nakalibrovat“ metodu tak, aby byla pozorována dobrá korelace mezi experimentálními a teoretickými výsledky. Vzhledem k tomu, že nejsou experimentální výsledky dostupné, byla metoda provedena bez možnosti validace s experimentálními výsledky a je nutné její výsledky vnímat v kontextu celé práce, tedy jako mezikrok výběru nejlepšího modelu komplexu a způsob zpřesnění výběru modelů získaných dokováním.

5.2.5.1 MM/PBSA výsledky a kontrola redundantních modelů

Pro každou variantu byly vyneseny střední hodnoty vazebných energií a standardní chyby. Na základě těchto výsledků bylo vybráno vždy 1-3 nejlepších modelů o nejvíce záporné vazebné Gibbsovo energii pro každý Myomedinový protein a pro MPER ligand. Sumarizace vybraných variant včetně jejich odhadnutých vazebných energií jsou uvedeny v **tabulce 16**. Zároveň byly vybrány oba případy, ve kterých dochází k vazbě předpokládaným smyčkovým módem – tedy modely 2 a 3 varianty MLA158. Výsledky pro všechny analyzované trajektorie modelů jsou zobrazeny v **přílohách B 6 a B 7**.

Pro každý z vybraných byla provedena kontrola redundance. Té bylo dosaženo tak, že byla pomocí balíku `gRINN` z MD trajektorie vypočtena matice intermolekulových interakcí (detailněji budou výsledky a postup budou diskutovány v **kapitole 5.2.8**). Kombinace každých dvou modelů v rámci jedné proteinové varianty byly následně analyzovány pomocí vlastního Pythonovského skriptu. Do tohoto skriptu vstupují matice párových intermolekulových interakcí. Matice jsou z povahy problému diagonálně symetrické a mají stejné rozměry, protože porovnáváme rozdílné konformace komplexů

identického složení. Výsledné matice jsou dále nazývány „diferenčními maticemi“. Algoritmus této metody je ve zkratce následující:

- a) Obě matice párových intermolekulových interakcí jsou standardizovány na hodnoty -1 až +1.
- b) Každá matice obsahuje majoritní počet párů reziduí neúčastnících se vazby, které tvoří „baseline“ hodnotu, která je nenulová. Tyto baseline hodnoty by činily následnou analýzou nemožnou (představme si, že na pozici XY matice 1 existuje interagující pár reziduí s volnou energií 10 kcal/mol (před normalizací); na identické pozici XY matice 2 žádný takový pár není, ale baseline hodnota je velmi malé číslo – výsledkem porovnání je velmi vysoké číslo, které efektivně „vtiskne“ matici 1 do výsledné diferenční matice). Tyto hodnoty by interferovaly se stanovením podobností dvou matic a jsou odstraněny. Toho je dosaženo vypočtením modu hodnot matice. Tak byla získána nejčastěji se objevující hodnota matice odpovídající baseline.
- c) Na některých místech hodnota baseline mírně osciluje okolo modové hodnoty, k této hodnotě je tak připočtena nízká hodnota taková, že jsou ideálně všechny informace nenesoucí hodnoty menší, než tato baseline hodnota.
- d) Všechny hodnoty nižší, než baseline hodnota jsou nahrazeny nulou.
- e) Pro každou pozici (x,y) matic 1 a 2 je provedena operace $\delta_{(x,y)_{1,2}} = |(x,y)_1 / (x,y)_2|$. Pokud je hodnota $(x,y)_1 > (x,y)_2$, je na pozici $(x,y)_{delta}$ diferenční matice vložena hodnota $\delta_{(x,y)_{1,2}}^{-1}$, výsledkem je tedy rozdíl dvou matic na pozici (x,y) v rozsahu 0-1. Pokud platí $(x,y)_1 < (x,y)_2$, je na pozici $(x,y)_{delta}$ vložena hodnota $\delta_{(x,y)_{1,2}}$. Pokud platí $(x,y)_2 = 0 \mid (x,y)_1 = 0$, je $\delta_{(x,y)_{1,2}} = 0$ (nulová podobnost – převážně kvůli důsledkům na vizualizaci a s ohledem na možné dělení nulou)
- f) Přehlednost vizualizace je dále zvýšena tak, že jsou z výsledné diferenční matice odstraněny hodnoty nižší, než 0.001 (nastavitelný parametr).

Tento postup umožňuje kvantitativně porovnávat rozdíly mezi vazbami odlišných konformací komplexů stejného složení s mnohem vyšší mírou jistoty, než pouhou vizuální inspekci výsledných struktur nebo hodnot energií vazeb. Výsledkem je „heatmap“ vizualizace diferenční matice, kde jsou vidět páry reziduí, které jsou pro interakce společné (jak s negativním, tak pozitivním příspěvkem) a jejich hodnota umožňuje kvantitativní pohled na interakci.

Získané diferenční matice sloužily jako kontrola redundancí modelů – pokud by se dva modely ukázaly jako výrazně podobné, byl by z daných modelů vybrán pro další analýzu jen jediný. Princip těchto matic je ilustrován v pro modely 10E8fr-MLA 158 2 a 3 v přílohách B 8 až B 10.

Tabulka 16 – Výsledky několika nejlepších MM/PBSA na dokování získaných modelů komplexů MLA-10E8 fragment. * jsou označeny modely kde je interakce řízena smyčkovým vazebným módem.

Model komplexu	$\Delta G_{\text{vaz,MM/PBSA}} / \text{kcal} \cdot \text{mol}^{-1}$; standardní chyba
MLA24 mod. 2	$-57,2 \pm 6,94$
MLA24 mod. 3	$-53,8 \pm 4,52$
MLA24 mod. 6	$-53,3 \pm 4,70$
MLA25 mod. 2	$-64,1 \pm 6,21$
MLA25 mod. 5	$-58,2 \pm 6,65$
MLA92 mod. 0	$-69,6 \pm 7,49$
MLA92 mod. 1	$-66,5 \pm 5,93$
MLA158 mod. 2*	$-43,5 \pm 6,02$
MLA158 mod. 3*	$-41,5 \pm 4,46$
MLA158 mod. 8	$-73,4 \pm 7,54$
WT mod. 8	$-78,5 \pm 6,33$
WT mod. 2	$-71,9 \pm 5,64$
WT mod. 6	$-70,8 \pm 7,53$
MPER mod. 2	$-75,5 \pm 5,02$
MPER mod. 6	$-66,5 \pm 4,50$
MPER nativní struktura	$-67,7 \pm 4,81$

Překvapivým výsledkem této analýzy bylo, že modely mající β -vazebný charakter dominují i v rámci MM/PBSA odhadnutých vazebných energiích. U všech variant byly nalezeny stabilní interakce s negativní změnou volné Gibbsovy energie při navázání. Síla vazby MyoWT byla nečekaně větší, než síla vazeb ostatních myomedinových proteinů. Jejich absolutní velikost nehraje v tomto kroku velikou roli – jak bylo řečeno, MM/PBSA je metoda vhodná k vzájemnému porovnávání modelů, nikoliv ke kvantitativním závěrům. Metoda samotná velmi nadhodnocuje vazebné energie. Tento zdánlivý nesoulad bude potvrzen a vysvětlen dále – nyní se spokojíme s vysvětlením, že nejspíše opravdu existuje možná nezanedbatelně silná vazba MyoWT na 10E8 (nulové hodnoty nebyly pozorovány ani v ELISA experimentech, viz. **kapitola 5.1.6**), nicméně nemá mimikující schopnosti a není biologicky relevantní.

5.2.5.2 MPER kontrola

Jako kontrola celého procesu sloužilo porovnání teoretických výsledků s experimentálně známými parametry vazby 10E8 Fab s nativním MPER ligandem. V části dokování a výběru modelů pomocí MM/PBSA by za správnou byla kontrola považována, pokud by kombinace dokování a MM/PBSA analýzy byla schopna vytvořit a odlišit takový model vazby, který by byl blízko vazbě z vyřešené struktury (která však nemusí popisovat vazbu v roztoku). Podle seřazených hodnot vazebné volné energie by byly aplikovaným postupem pro další analýzu vybrány modely komplexů 10E8fr.-MPER 2 a 6 (-66,5 a -75,5 kcal/mol). Porovnáním s hodnotou zjištěnou MM/PBSA analýzou pro vyřešenou strukturu (-67,7 kcal/mol) je vidět, že jsou blízko oba vybrané modely. Detailnější kontrola proběhla porovnáváním RMSD hodnot mezi vyřešenou strukturou a modely. Nutné je mít na paměti, že jsou struktury flexibilní a toto porovnání zahrnuje pouze jednu strukturu (tedy získanou krystalografií a rigidním dokováním). Přesnější porovnání pak bylo provedeno na základě párových interakcí v **kapitole 5.2.8**.

Superimpozice všech tří struktur na sebe je v různých reprezentacích a pohledech zobrazena na **obrázku 24**. Patrné je, že je v shodě nativní s dokovanými a vybranými strukturami vyšší v případě druhého α -helixu a nižší v případě prvního. Rozdíly byly tedy analyzovány pro dvě jednotlivé části MPER peptidu – 1. helix (NEQELLELDKWASL) a druhý helix (WFDITNWLWYIRR). Dle původního článku o vyřešené struktuře

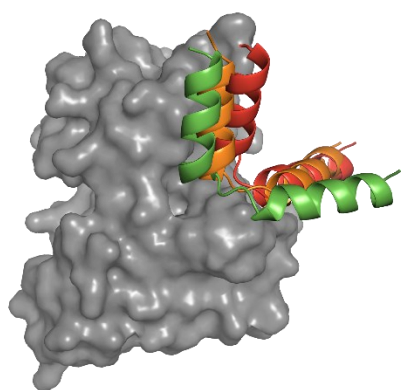
10E8Fab-MPER fragmentu se klíčová rezidua pro vazbu nachází ve spojce a druhém helixu MPERu (v 4G6F rezidua N671, W672, F673, T676, R682).⁸⁸ S ohledem na tuto skutečnost byla pozornost věnována zejména porovnání RMSD těchto reziduí mezi vyřešenou strukturou a modely.

RMSD vzdálenosti (vypočteno pomocí UCSF Chimera) hlavních řetězců a klíčových reziduí po superimpozici protilátkových fragmentů na sebe jsou uvedeny v **tabulce 17**. Je patrné, že je model 6 celkově blíže vyřešené krystalové struktuře. V obou modelech je druhý helix (klíčový helix) blíže krystalové konformaci vůči fragmentu než helix první. V 10E8fr.-MPER modelu 6 je hodnota RMSD nejlepší a poměrně blízko běžně udávanému „limitu kvality dokování“ 2 Å.

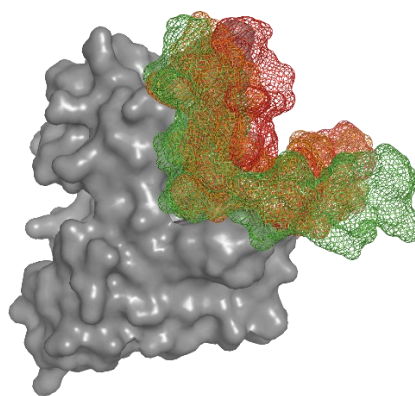
Tabulka 17 – RMSD vzdálenosti hlavních řetězců dvou predikovaných konformací vazby ligandu MPER na protilátkový fragment 10E8.

Část	RMSD ke známé struktuře / Å
Helix 1, model 2	3,079
Helix 2, model 2	4,573
Helix 1, model 6	12,443
Helix 2, model 6	3,049
N671, model 2	3,825
W672, model 2	4,643
F673, model 2	5,126
T676, model 2	4,191
R682, model 2	5,333
N671, model 6	2,484
W672, model 6	1,372
F673, model 6	1,722
T676, model 6	2,282
R682, model 6	2,048

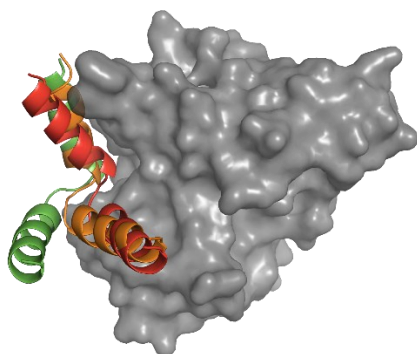
Kombinace dokování a MM/PBSA selekce byla tedy schopna vygenerovat a odlišit v případě komplexu 10E8fr.-MPER strukturu blízko krystalové. Obě vybrané struktury jsou v případě klíčového interagujícího 2. helixu blízko vyřešené struktuře. Méně důležitý 1. helix je v případě modelu 6 výrazně vychýlen oproti vyřešené struktuře.



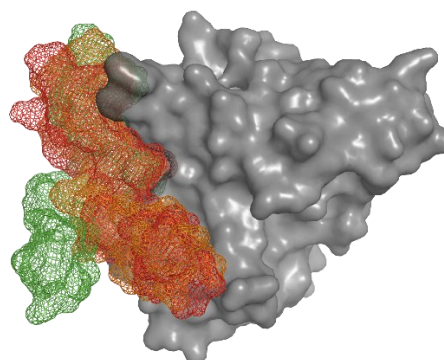
2°-struktura, pohled zepředu



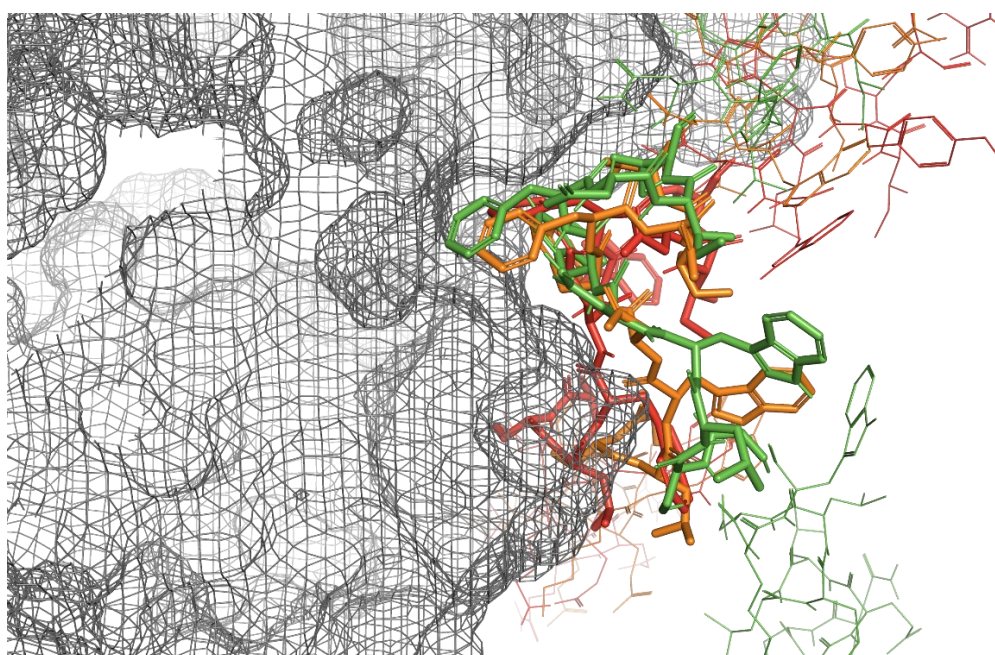
Povrch, pohled zepředu



2°-struktura, pohled zezadu

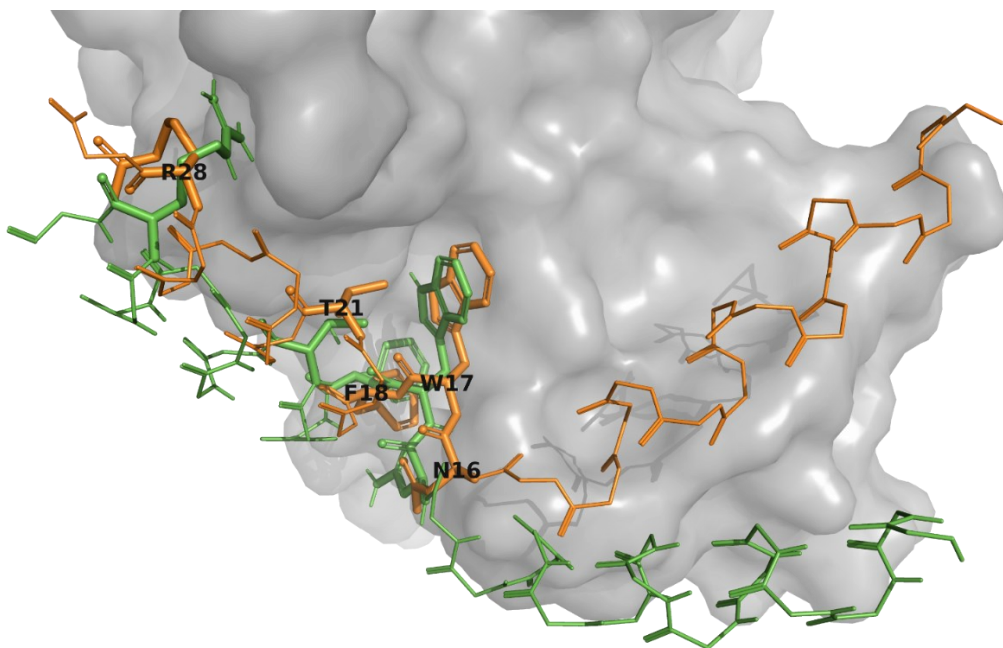


Povrch, pohled zezadu



Detail na kontakty, tlustě jsou rezidua v „interakční kapse“.

Obrázek 24 – Porovnání nativní a predikované struktury protilátkového fragmentu 10E8 s MPER ligandem. Oranžově je zobrazena nativní konformace, červeně model 2, zeleně model 6. Šedě je zobrazen protilátkový fragment.



Obrázek 25 – Klíčová interakční rezidua vyřešené krystalové struktury komplexu 10E8fr.-MPER a modelu 6 z dokování. Oranžově je zobrazena vyřešená struktura, zeleně model 6. Slabě je zobrazen hlavní řetězec MPER peptidu, tlustě jsou uvedeny klíčové postranní řetězce pro interakci.

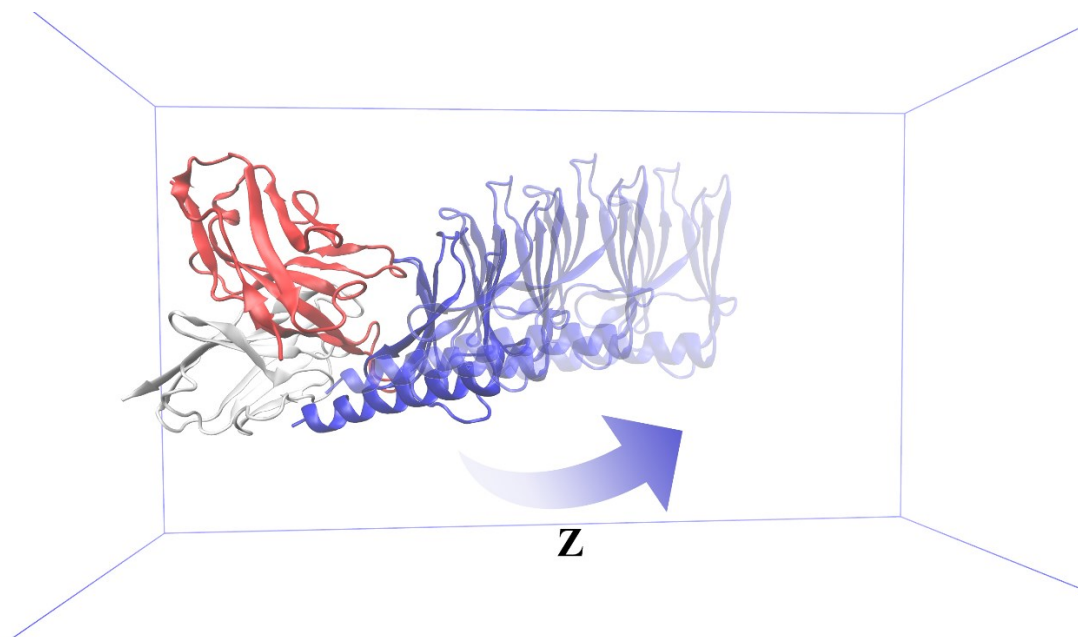
5.2.6 Coarse-grained Umbrella Sampling

Dalším krokem byla přesnější analýza protein-protein vazby protilátkového fragmentu s variantami MLA, MyoWT a nativním ligandem MPER. Pro tento účel byla zvolena metoda „tahání“ jednoho z proteinů od druhého (po ose z) do vzdálenosti 10 nm od těžišť s využitím harmonického potenciálu a následné umbrella sampling simulace. Znázornění trajektorie tahání ukazuje **obrázek 26**. Samotné analýzy pak proběhly pouze na prvních 5 nm těchto trajektorií, během kterých bylo uspokojivě dosaženo ztrátě všech intermolekulových kontaktů.

Analýza výsledných sil působících na pomyslnou pružinu proběhla metodou WHAM, která vyústila v hodnoty PMF („potential of mean force“) systému v průběhu trajektorie tahání. Rozdíl mezi minimem a maximem těchto průběhů byl uvažován jako ΔG_{vazba} proteinu/peptidu k protilátkovému fragmentu.

Pro přípravu trajektorie („pulling“/tahání) byly všechny proteiny nejdříve umístěny do boxu s identickými vektory X a Y (dle velikosti systému tak, aby byl odstup proteinu

alespoň 2.0 nm) a s velikostí vektoru Z rovnou 17 nm. Proběhla solvatace, energetická minimalizace, ekvilibrace a nakonec samotné tahání, jak je uvedeno v kapitole 4.2.3.4.



Obrázek 26 – Trajektorie tahání MLA proteinu od protilátkového fragmentu ve směru osy Z . Modře jsou vyznačeny hrany simulačního boxu. Bílo-červeně je vyznačena sekundární struktura protilátkového fragmentu, modře (se stoupající průhledností) některé po sobě jdoucí kroky tahání.

Výše uvedený postup platí pro tento i další krok (US s atomistickým rozlišením). Pro zrychlení celého procesu bylo nejprve provedeno první kolo s pomocí „coarse-grained“ aproximace pod silovým polem SIRAH. Podobný postup byl popsán například autory Patelem a Ytrebergem v roce 2018, kde byla prokázána jeho relativně vysoká přesnost.⁹⁴

Po vygenerování průběhů PMF byly vybrány nejlepší modely. Vyhodnocení proběhlo prakticky tak, že byl průběh PMF omezen na trajektorii 5 nm (dostatečná k úplné ztrátě intermolekulových kontaktů a úplné solvataci obou partnerů) a její maximum bylo odečteno. Minimum PMF trajektorie bylo uvažováno jako navázaný stav, následně byl odečten rozdíl mezi těmito dvěma extrémy. Seznam vybraných modelů je uveden v **tabulce 18**. Díky četným obtížím se ani přes opakované pokusy nepodařilo touto metodou stanovit volnou energii vazby s MPER ligandy, které se během simulací deformovaly a neposkytovaly „rozumné“ průběhy PMF. Ve všech případech však byla (alespoň zdánlivě) hodnota Δ PMF vyšší pro model 6, než pro model 2. I z tohoto důvodu

(a z důvodu lepšího RMSD oproti krystalové struktuře), byl vybrán pro další analýzy právě tento model.

Tabulka 18 – Nejlepší modely komplexů 10E8 fragment – MLA protein vybrané pomocí coarse-grained umbrella sampling simulací. Hodnoty ΔPMF odpovídají změně Gibbsovy volné energie při tahání vazebného proteinu po směru osy Z od protilátkového fragmentu. Uvedeny jsou i tři konformace komplexu s nativním ligandem MPER, který byl využit jako kontrola celého procesu.

Model / 10E8 fragment	$\Delta\text{PMF} / \text{kJ} \cdot \text{mol}^{-1}$
MLA024 mod. 6	46,25
MLA025 mod. 5	58,66
MLA092 mod. 0	63,73
MLA158 mod. 2	45,12
MLA158 mod. 3	62,3
MLA158 mod. 8	66,96
MyoWT mod. 6	84,09
MyoWT mod. 8	67,12
MPER mod. 2	x
MPER mod. 6	Vybrán na základě RMSD
MPER vyřešená struktura	x

5.2.7 Umbrella Sampling s atomistickým rozlišením

Obdobnou metodou, jaká byla popsána v předchozí kapitole (5.2.6) byly provedeny i simulace a analýzy tří modelů vazby 10E8fr.-MLA158 – modely 8 (β -vazebný) a modely 2+3 (vzájící oblastmi aminokyselinových záměn). Tentokrát však proběhly simulace s použitím atomistického popisu silovým polem CHARMM36m se všemi kroky (kromě tahání a samotného US) obdobnými jako v simulacích popsaných **kap. 5.2.3**. Výsledné křivky průběhů PMF (pouze do 5 nm vzdálenosti těžišť molekul) podél trajektorie tahání jsou zobrazeny v **přílohách B 11 až B 13**.

Zjištěné Δ PMF (ekvivalentní Δ G simulované trajektorie napodobující vazbu) pro modely vazby 10E8fr.-MLA158 byly 33,34 kJ/mol (mod. 2), 38,42 kJ/mol (mod. 3) a 44,97 kJ/mol (mod. 8). Pořadí síly vazeb zde souhlasí s CG US simulacemi, hodnoty jsou však celkově nižší (CG US simulace vyústily v 45,12; 62,3; 66,96 kJ/mol pro mod. 2; mod. 3 a mod 8). Z povahy atomistického popisu molekul těchto simulací (oproti zjednodušenému CG popisu) by měly být zde zmíněné výsledky přesnější.

5.2.8 Analýza intermolekulových interakcí komplexů s MPER

Analýza párových interakcí proběhla pomocí aplikace gRINN s využitím posledních 5 ns (každý snímek) MD simulací vybraných modelů komplexů 10E8fr. – MPER/MLA a pomocí vyřešené struktury komplexu s MPER ligandem. Posledních 5 ns bylo vybráno univerzálně pro všechny analýzy všech komplexů (MPER i MLA) kvůli dosažení konvergence ukazatelů (RMSD, gyrační poloměr) kvality MD simulací, jak je vidět na **přílohách B 4 a B 5**.

Dále uvedené hodnoty interakčních volných energií pochází z průměrů volných energií v průběhu snímků 5 ns trajektorie, provedené pod silovým polem CHARMM36m. Zjištěny byly tedy s využitím přibližných potenciálů a jejich hodnoty nemusí přesně reflektovat realitu.

Dle alaninového skenování provedeném autory pdb:4G6F struktury (náhradou aminokyselin za alanin a pozorováním vazebné afinity MPER-10E8 Fab) bylo vybráno několik pro nativní vazbu klíčových reziduí protilátky 10E8, umístěné ve dvou distinktních oblastech na těžkém řetězci:⁹²

a) CDR H3 region (oblast 1/site 1/S1): Y99, F100a, G100d, P100f, P100g, E100i, E100j

b) Hydrofobní štěrbina – W33, R50, E53. Označována HC (“hydrophobic cleft”)

Z hlediska MPER ligandu byla identifikována 4 klíčová rezidua, vážící oba regiony 10E8 Fab.⁹²

A) W680, R/K683 – CDR H3 region (oblast 1)

B) W672, F673 – hydrofobní štěrbina

Na základě teorie o mimikování interakcí MPER s 10E8 protilátkou pomocí MLA proteinů byly tyto oblasti 10E8 analyzovány pro všechny vybrané modely na síly, charakter a stability interakcí pomocí matic párových interakcí získaných aplikací gRINN.

Nejdůležitější interakční místa zjištěná experimentálně u vyřešené struktury komplexu v porovnání s dokováním získaným modelem 6 jsou vizualizována na **obrázku 28**. Patrné je, že v krystalové struktuře komplexu interagují rezidua MPERu Trp25 a Arg28 s oblastí 1 protilátky (tvořenou Tyr99, Phe100a, Trp100b). Do hydrofobní štěrbiny protilátky (tvořené primárně Trp33 a Arg50) zapadají Trp17 a Phe18 MPERu. Interakce je vedena primárně druhým helixem (na obrázku nahoře).

5.2.8.1 Vazba MPER vyřešené struktury a modelu 6 z pohledu všech reziduí

Vizuální podoba matice volných energií dle párových interakcí pro posledních 5 ns MD trajektorie komplexu 10E8fr.-MPER (pdb:4G6F – vyřešená struktura) je zobrazena b **příloze B 14** (pro model 6 pak v **příloze B 15**). V následujícím textu jsou řetězce označeny A (těžký řetězec protilátky), B (lehký řetězec protilátky) a C (MPER peptid). Tyto označení jsou uvedeny před aminokyselinami jako identifikátor jejich pozice ve struktuře a kterému partneru náleží.

5.2.8.1.1 Vyřešená struktura 10E8 fr.-MPER

Z hlediska nejsilnějších párových interakcí vedou významně interakce ALys97-CAsp664 (-75,7 kcal/mol), AAsp100-CArg683 (-17,4 kcal/mol), AGly100d-CArg683 (-16,1 kcal/mol), AGlu53-CTrp672 (-11,3 kcal/mol). Interakce jsou převážně facilitované těžkým řetězcem protilátky. První helix MPER nese nejsilnější interakci (skrze CAsp664), zbytek ze 4 nejsilnějších interakcí je nesen druhým helixem.

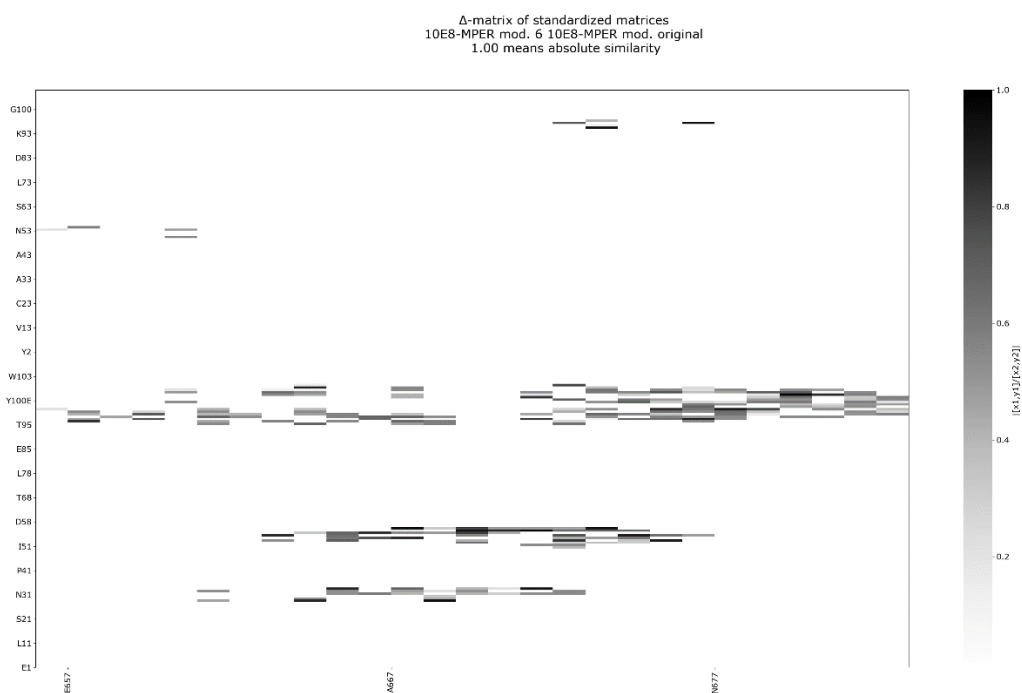
Interakce ALys97-CAsp664 je enormně silná. Mohlo by se jednat o artefakt stanovení a nebo použitého silového pole, či o opravdu nesmírně silnou iontovou (s možnou dodatečnou vodíkovou) interakci s aminokyselinovými zbytky zabořenými za nepřístupnou rozpouštědla (v prostředí s nízkou permivitou ϵ). Kontrolu solvatace interagujících

molekul by bylo možné provést v samotném balíku GROMACs, v této práci však nebyla provedena.

5.2.8.1.2 10E8 fr. – MPER model 6

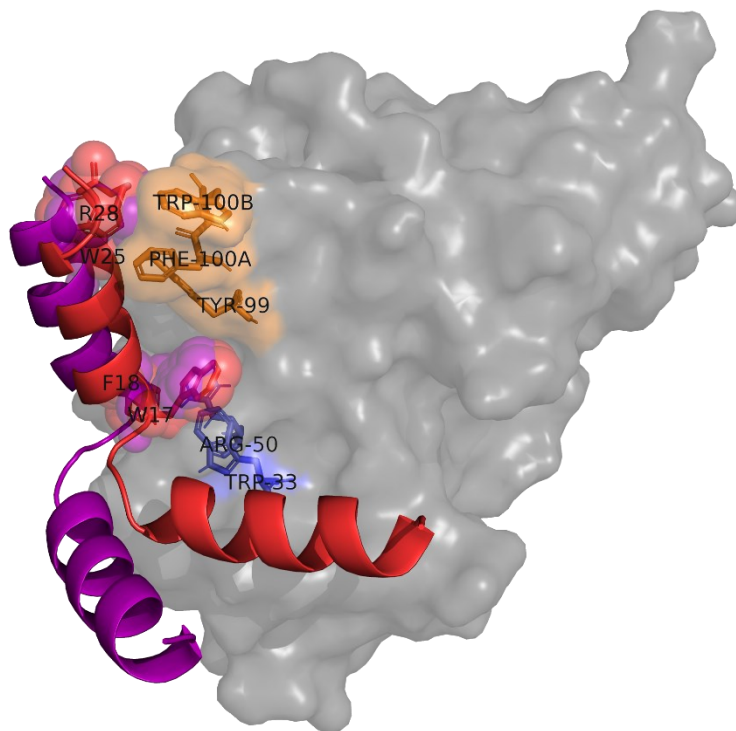
Nejsilnější párová interakce je ALys97-CAsp664 (-31,3 kcal/mol). Dále jsou nejsilnější interakce AGlu53-CTrp672 (-14,4 kcal/mol), AArg94-CAsp664 (-9,90 kcal/mol) a AGlu53-CAsn671 (-8,00 kcal/mol).

Na základě těchto čtyř nejsilnějších stabilizujících interakcí je patrné, že je model získaný dokováním blízko vyřešené struktury komplexu. Zachována je nejsilnější interakce (ALys97-CAsp664), byť v oslabené podobě modelu 6 oproti originální struktury (být je hodnota pro vyřešenou strukturu, -75,7 kcal/mol extrémně vysoká a mohla by být pouze artefaktem). Zachována je také interakce AGlu53-CTrp672, zbylé dvě ze čtyř nejsilnějších interakcí jsou pro model a vyřešenou strukturu rozdílné.



Obrázek 27 – Porovnání párových interakcí mezi vyřešenou strukturou komplexu 10E8 fr. – MPER (pdb:4G6F) a modelem 6 získaným dokováním dvou partnerů. Tmavší barva označuje vyšší podobnost v dané párové interakci.

Pro vizuální porovnání vazebných módů slouží matice zobrazená na **obrázku 27**. Získána byla porovnáním matic párových interakcí mezi oběma strukturami. Vidět je podobnost obou modelů. Poměrně příznivé výsledky RMSD klíčových reziduí i celku mezi strukturami modelu 6 a vyřešenou strukturou již byly diskutovány v **kap. 5.2.5.2**.



Obrázek 28 – Nejdůležitější experimentálně stanovená interagující rezidua upraveného komplexu pdb:4G6F (10E8 fr. – MPER). Červeně je vyznačen peptid MPER (vyřešená struktura) a jeho sekundární struktura. Obdobně je fialově vyznačen MPER model 6. Jednopísmenným kódem, pozicí, tyčkami a průhlednými koulemi jsou označena nejdůležitější rezidua (určená experimentálně). Ekvivalentně jsou oranžově/modře (oblast 1 interakce/hydrofobní štěrbin) označeny nejdůležitější pozice 10E8 fragmentu. Šedě je zobrazen povrch protilátkového fragmentu.

5.2.8.2 Vazba MPER vyřešené struktury a modelu 6 z pohledu mimikujících reziduí

5.2.8.2.1 Vyřešená struktura MPER

Celková síla párových interakcí trajektorie krystalové struktury 10E8fr.-MPER činila -82,45 kcal/mol (-49,74 elstat; -32,71 vdW). Z toho -21,38 kcal/mol (-15,67 elstat. ; -5,71 vdW) připadá na interakce MPERu s hydrofobní štěrbinou (HC) protilátkového fragmentu (zejména s AGlu53). -61,07 kcal/mol (-34,07 elstat; -27,00 vdW) přitom připadá na interakce s oblastí 1 (S1).

Z hlediska reziduí protilátky jsou tak interakce MPERu nejsilnější s AGlu53 hydrofobní štěrbinou (převážně elektrostatické interakce s trojicí reziduí CAsn671-CTrp672-CPhe673 o volných energiích -7,13 ; -11,3 ; -5,38 kcal/mol). Stabilizující je i interakce MPER s ATrp33 (převážně van der Waalsovské interakce) o volné energii -3,22 kcal/mol s nejsilnějším partnerem MPER CTrp672 (-1,63 kcal/mol, z toho -1,92 kcal/mol vdW charakteru s drobnou elektrostatickou repulzí). Existují i relativně mírně destabilizující interakce MPER s AArg50 o celkové síle 3,25 kcal/mol (převážně elektrostatické repulze, zejména pocházející od CTrp672). Přehled volných energií párových interakcí zobrazuje **tabulka 19**.

Z oblasti 1 protilátky interagují zejména rezidua ATyr99-APhe100A-AGly100D-APro100F-APro100G (-10,5 ; -10,9; -17,0; -9,32; -8,98 kcal/mol). Nejzávažnější jsou tyto interakce:

- A) Interakce s ATyr99 je téměř úplně vdW charakteru a je facilitována primárně CLeu679, CTrp680, CThr676 a CTrp672 (-2,50 ; -2,35; -2,10; -1,07 kcal/mol)
- B) Interakce s APhe100a je přibližně stejně charakteru elektrostatického i van der Waalsovského. Nejsilnější stabilizující interakce je s CArg683 MPER o volné energii -9,26 kcal/mol (z toho -8,83 kcal/mol elektrostatických), hydrofobní interakce pochází z řady slaběji interagujících reziduí, zejména druhého helixu MPER.
- C) AGly100d poskytuje s MPER vůbec nejsilnější interakce celého komplexu (-17,0 kcal/mol; z toho -16,5 elstat. charakteru). Velice silná je interakce s CArg683

(-16,1 kcal/mol, elstat. charakter), nezanedbatelná je i z větší části elektrostatická interakce s CTrp680 (-2,35 kcal/mol z toho 1,84 kcal/mol elstat.)

D) Dvojice prolinů APro100f a APro100g poskytuje -9,32 ; -8,98 kcal/mol volné energie. APro100f i APro100g interagují s Trp672 (-0,56; -5,09 kcal/mol obojí převážně elstat.), Phe673 (-2,33 ; -1,71 kcal/mol převážně vdW). Výrazné jsou i interakce s CThr676 (-3,00 ; -2,14 kcal/mol, spíše hydrofobní ale s nezanedbatelnou elstat. komponentou) a dalšími. APro100f má stabilizující interakce s CAsn667 (-1,56 kcal/mol) a s CTrp680 (-1,85 kcal/mol).

Z hlediska trajektorie vyřešené struktury MPER ligandu jsou nejdůležitější stabilizující interakce komplexu s 10E8fr. nesené rezidui CArg683 (ob. 1); CTrp672 (HC+ob.1); CPhe673 (HC+ob.1); CThr676 (přev. ob.1); CAsn671 (přev. HC); CTrp680 (ob.1) (-25,70; -20,01; -9,85; -8,04; -7,80; -6,47 kcal/mol).

Rezidua 10E8fr. byla vybrána na základě uvedeného zdroje a tak není možno důležitost těchto reziduí nijak validovat použitými metodami. Oproti tomu je možné zkorelovat experimentální výsledky důležitých reziduí MPER s výsledky získanými analýzou párových volných vazebných energií. Nejdůležitější rezidua pro vazbu 10E8fr. na ligandu MPER (CArg683, CTrp672, CPhe673, CTrp680) byla z hlediska volných energií nejdůležitějšími rezidui v provedené analýze. Na základě analýzy volných energií párových interakcí byla zároveň identifikována i další dvě energeticky důležitá rezidua MPER – CThr676 a CAsn671.

Předpokladem autora je, že pro úspěšné mimikování povrchu MPER ligandu pomocí MLA proteinů by měly být interakce podobné zde popsaným (tedy ideálně podobná rezidua interagující se stejnými místy 10E8). Pro každý analyzovaný model bude tedy provedeno porovnání. Nejprve však bude provedeno porovnání s modelovanou vazbou MPER na 10E8 fragment, vybranou pomocí MM/PBSA metody.

5.2.8.2.2 MPER model 6 – kontrola dokování pomocí zachování interakcí

Tabulka 19 – Přehled volných energií párových interakcí experimentálně identifikovaných důležitých reziduí 10E8 s MPER ligandem z krystalové struktury v porovnání s modelem MPER mod. 6. Červeně jsou vyznačeny stabilizující, modře destabilizující interakce. Hodnoty jsou zaokrouhleny na 4 platné číslice.

Kcal/mol	TRP33	ARG50	GLU53	TYR99	PHE100A	GLY100D	PRO100F	PRO100G	GLU100I	GLU100J	Σ
MPER kryst. Σ	-3.216	3.246	-21.41	-10.54	-10.87	-16.95	-9.321	-8.976	-1.251	-3.158	-82.45
MPER kryst. elstat.	-0.5436	3.360	-18.48	-0.8754	-5.974	-16.45	-1.534	-5.286	-0.9279	-3.025	-49.74
MPER kryst. vdW	-2.672	-0.1143	-2.925	-9.665	-4.900	-0.5029	-7.786	-3.690	-0.3227	-0.1337	-32.71
MPER mod. 6 Σ	-3.884	2.072	-13.87	-8.035	-13.03	-5.751	-5.322		-2.245		-57.63
MPER mod. 6 elstat.	-0.653	2.904	-8.650	1.666	-6.499	-3.943	-2.038		-1.888		-19.10
MPER mod. 6 vdW	-3.230	-0.8318	-5.224	-9.701	-6.528	-1.807	-3.285		-0.35679		-30.96

Poslední porovnání modelu 6 komplexu 10E8fr.-MPER s vyřešenou strukturou proběhlo z hlediska energeticky důležitých párových interakcí. Celková síla párových interakcí činila -50,06 kcal/mol (-elstat. -19,10; vdW. -30,96), oproti celkovým -82,45 kcal/mol vyřešené struktury, kde je však nejsilnější interakce (ALys97-CAsp664) „podezřele“ silná a mohla by být výsledkem artefaktu. Oproti vyřešené struktuře je celkový charakter stabilizujících interakcí spíše van der Waalsovský.

Interakce s hydrofobní štěrbinou poskytují -15,67 kcal/mol (oproti -21,38 vyř. str.) a jsou o něco více hydrofobního charakteru (-9,29 kcal/mol) než u vyř. str. (-5,71 kcal/mol). S oblastí 1 poskytují interakce dohromady -34,38 kcal/mol (-21,68 vdW) oproti -61,07 kcal/mol (-27,00 vdW) u vyř. str. Poměrově je tedy i interakce modelu 6 s oblastí 1 více van der Waalsovského charakteru.

Nevíce stabilizující interakce v oblasti hydrofobní štěrbině je s AGlu53 (-13,87 kcal/mol, z toho -5,22 kcal/mol vdW) pomocí reziduí CAsn671-CTrp672-CPhe673 (-8,02; -14,44; -3,51 kcal/mol; převážně elektrostatického charakteru). Tyto interakce jsou tak zachovány (a v případě CAsn671-CTrp672 dokonce zesíleny) v porovnání s vyř. str. Zachována je i vdW interakce CTrp672 s ATrp33, stejně jako mírně destabilizující

interakce CTrp672 s AArg50. Z hlediska interakcí s hydrofobní štěrbinou je model 6 konzistentní s vyřešenou strukturou.

Z oblasti 1 protilátky interagují u modelu 6 zejména rezidua ATyr99-APhe100A-AGly100D-APro100F (hodnoty interakcí pro residuum APro100G se nepodařilo stanovit). Celkové síly těchto interakcí těchto reziduí jsou -8,03; -13,03; -5,75; -5,32 kcal/mol.

- A) Interakce s ATyr99 je téměř úplně vdW charakteru a je facilitována primárně CThr676, CLeu679, CTrp680 a CTrp672 (-2,41; -2,35; -1,93; -1,46 kcal/mol), obdobně jako u vyřešené struktury.
- B) Interakce s APhe100a je zhruba stejně elstat. i vdW charakteru. Nejsilněji s tímto residuem interaguje CArg683 s volnou energií -6,16 kcal/mol, z toho -5,29 kcal/mol elstat. charakteru. Interakce s tímto residuem je tedy obdobná u obou struktur. Interakce s tímto residuem je u modelu 6 silnější než u vyř. str.
- C) Interakce s AGly100d je oproti vyř. str. oslabena (-5,75 vs. -16,95 kcal/mol). Interakce s CArg683 je zanedbatelná, nahrazuje ji do jisté míry interakce se sousedním CArg684 (-1,99 kcal/mol, elstat. charakter).
- D) APro100f má celkové interakce s MPER modelem 6 o volné energii -5,32 kcal/mol. Nejsilnější jsou interakce s CAsn677 a CThr676 (-1,56; -1,13 kcal/mol). Oproti vyř. struktuře je interakce odlišná, avšak stále stabilizující. Zachována je pouze oslabená interakce s CThr676.

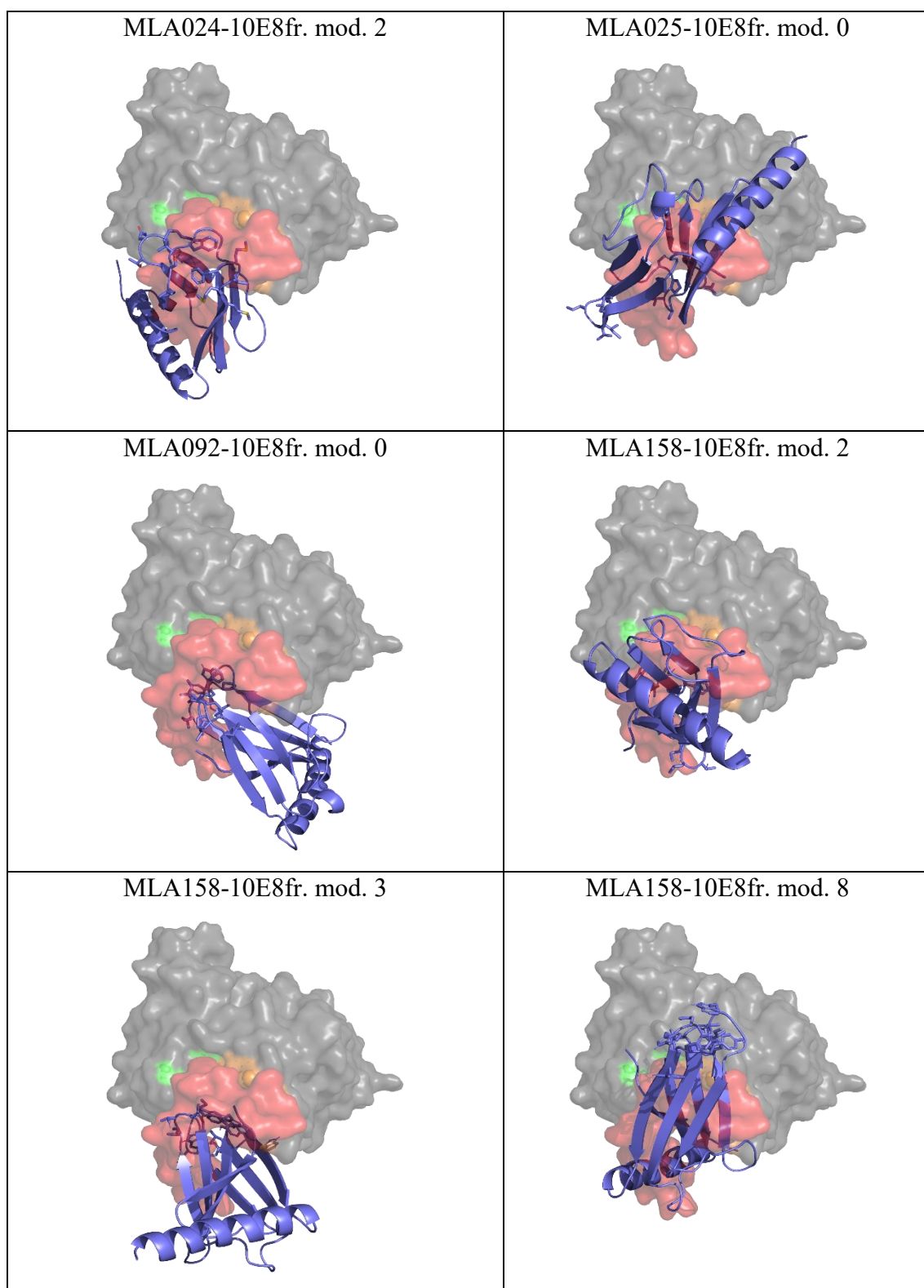
Na základě informací diskutovaných zde a v **kap. 5.2.5.2** je možné říct, že bylo dokováním a následným výběrem pomocí MM/PBSA nalézt model komplexu, který jak svojí geometrií (byť je první, méně důležitý helix vyosen od protilátky) tak volnými energiemi párových interakcí uspokojivý, vzhledem ke krystalové struktuře. Na paměti je nutné mít, že vyřešená krystalová struktura negarantuje získání nativní struktury. Ta je jednak dynamická, v roztoku na ní působí odlišné síly oproti krystalu a existuje v kontextu celého antigenu, což nebylo uváženo v experimentech ani krystalografických ani zde popsaných teoretických.

5.2.9 Analýza intermolekulových interakcí komplexů s MLA proteiny

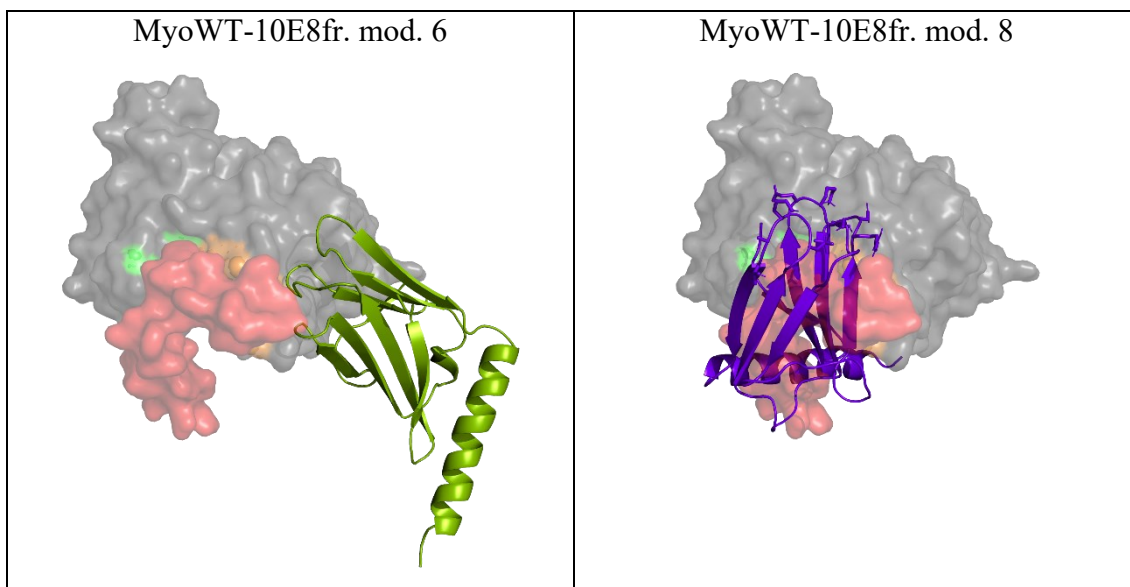
Přehled volných energií párových interakcí MLA (a MyoWT) proteinů s klíčovými rezidui protilátkového fragmentu 10E8fr. je uveden v **tabulce 20**. Vizualizovaná podoba interakcí je na **obrázích 29 a 30**.

Tabulka 20 - Přehled volných energií párových interakcí experimentálně identifikovaných důležitých reziduí 10E8 s MPER ligandem z krystalové struktury v porovnání s modely komplexů s MLA proteiny. Červeně jsou vyznačeny stabilizující, modře destabilizující interakce. Hodnoty jsou zaokrouhleny na 4 platné číslice.

<i>Kcal/mol</i>	TRP33	ARG50	GLU53	TYR99	PHE100A	GLY100D	PRO100F	GLU100I	GLU100J	Σ
MPER kryst. Σ	-3.216	3.246	-21.41	-10.54	-10.87	-16.95	-9.321	-1.251	-3.158	- 82.45
MLA024 mod. 2 Σ	1.111			-0.5536	-0.3003	-0.1217	1.783	0.2888		2.208
MLA025 mod. 5 Σ	-0.2905		11.55	-14.98	-21.05	1.002	-3.198	-0.04301		- 27.00
MLA092 mod. 0 Σ	0.1519		-8.461	-5.663	-21.13	-2.052	0.7702	0.1763		- 36.21
MLA158 mod. 2 Σ	-0.3595		2.444	-6.885	-6.736	-1.603	1.136	-1.256		- 13.26
MLA158 mod.3 Σ	-6.037	0.04654	-4.849	- 0.06955	-0.9727	0.08108	-0.1969	-0.1083		- 12.11
MLA158 mod. 8 Σ	0.2283	0.2397	-0.5165	-14.86	-20.18	0.05233	-0.9643	2.292		- 33,72
MyoWT mod. 6 Σ				1.245	-5.783	-0.3916	-1.621	-81.77		- 88,32
MyoWT mod. 8 Σ	- 0.04091		10.59	-23.23	-9.875	2.984	3.356	8.364		-7,85



Obrázek 29 – Struktura vybraných komplexů MLA-10E8fr. analyzovaných z hlediska párových interakcí. Šedě je zobrazen povrch fragmentu 10E8 protilátky, zeleně/oranžově jsou na něm vyznačeny důležitá rezidua hydrofobní štěrbin/oblasti 1. Červeně je vyznačen povrch náležící MPER ligandu (z krystalové struktury) pro porovnání konformací. MLA proteiny mají vyznačenou sekundární strukturu a aminokyselinové záměny (tyčky).



Obrázek 30 – Struktura vybraných komplexů MLA-10E8fr. analyzovaných z hlediska párových interakcí. Šedě je zobrazen povrch fragmentu 10E8 protilátky, zeleně/oranžově jsou na něm vyznačeny důležitá rezidua hydrofobní štěrbin/oblasti 1. Červeně je vyznačen povrch náležící MPER ligandu (z krystalové struktury) pro porovnání konformací. MyoWT proteiny mají vyznačenou sekundární strukturu a aminokyselinové záměny (tyčky).

Vizualizované matice párových interakcí mezi MLA proteiny a 10E8 fragmentem jsou zobrazeny **přílohách B 16 až B 23**. Následovat bude popis interakcí jen důležitých reziduí, přibližný obraz o roli všech ostatních je možno získat ze zmíněných dodatků, přesný popis však není v této práci uveden.

5.2.9.1.1 MLA024 mod. 2

Na základě přehledu je zřejmé, že model komplexu 10E8fr./MLA024 (mod. 2) má bilanci interakcí s důležitými pozicemi 10E8 kladnou. Z hydrofobní štěrbiny interaguje pouze s ATrp33 (destabilizující), z oblasti 1 má pouze 3 zanedbatelně stabilizující interakce a dvě destabilizující. Usuzovat něco o mimikujících interakcích na základě těchto výsledků nelze. Z *in vivo* i *in vitro* experimentů je jisté, že protein MLA024 s 10E8 protilátkou interaguje (byť nejslaběji ze zkoumaných MLA variant).

Možná tři vysvětlení tohoto zdánlivého rozporu jsou následující:

- a) Skutečný vazebný mód je odlišný a nebyl v této práci identifikován. Toto vysvětlení zní pravděpodobně – nebyly analyzovány všechny získané modely komplexů pro tuto variantu, mezi kterými by mohl být ukrytý skutečný mód. Zároveň zde nebyly identifikovány žádné interakce skrze aminokyselinové záměny, což nahrává tomuto vysvětlení.
- b) Skutečný vazebný mód je identický a k mimikování stačí 3 slabé příznivé interakce s prvními třemi rezidui protilátkové oblasti 1.
- c) Zde provedené analýzy párových interakcí jsou nepřesné. Z povahy přijatelné přesnosti současných silových polí pro popis proteinů je názorem autora, že by se jednalo spíše o chybu experimentálního procesu než použitých aproximací.

U modelu 2 komplexu MLA024 nebyly žádné MPER kopírující interakce nalezeny. Aminokyselinové záměny zde s 10E8 vůbec neinteragují.

5.2.9.1.2 MLA025 mod. 5

Model 5 komplexu 10E8fr.-MLA025 má zápornou energetickou bilanci párových interakcí (celkově -27,0 kcal/mol; -8,19 elstat.; -18,9 vdW – z 30,33 % elektrostatické interakce). Jeho charakter vazeb je tedy oproti vyř. str. poměrově více vdW charakteru (MLA025 mod. 5 energeticky 69,67 % van der waalsovské interakce ; MPER 39,66 %). Z pohledu konkrétních reziduí protilátky je zde slabá stabilizující interakce s ATrp33 zejména skrze CAsp62 (-0,33 kcal/mol), kdežto v MPER s ATrp33 nejsilněji interaguje CTrp672. V MPERu hydrofobní Trp je tedy nahrazen nabitým Asp. Naopak výrazně destabilizující je interakce s AGlu53 (11,55 kcal/mol) způsobená primárně elektrostatickou repulzí s aminokyselinou CAsp62. Interakce s jinými rezidui hydrofobní štěrbinu nebyly zjištěny, jako celek tak β –mód vybraného modelu vazby MLA025 s 10E8 interaguje s hydrofobní štěrbinou výrazně energeticky nevýhodně.

Interakce s oblastí 1 je zde také oslabena oproti MPER – činí -38,3 kcal/mol (oproti MPER -61,1 kcal/mol). Interakce s reziduem ATyr99 (-15,0 kcal/mol) je silnější, než v MPER (-10,5) a je nesena zejména CGlu69 (ATyr99-CGlu69 -18,1 kcal/mol) a naopak destabilizována sousedním CGlu65 (4,41 kcal/mol) – obojí primárně elektrostaticky. Oproti této situaci má MPER interakce van der Waalsovské a nesené rezidui Leu, Trp a Thr.

Interakce s APhe100a je zhruba dvakrát silnější než u MPER (zde -21,1 kcal/mol, u MPER -10,9 kcal/mol). Může za ni zejména CArg28 (-13,2), tato situace je analogická situaci s MPER, kde CArg35, váže silně APhe100a.

Interakce s AGly100d protilátky jsou slabě destabilizující, zejména díky CGlu20 (3,48 kcal/mol), nejsilnější záporný příspěvek k volné energii poskytuje CArg28 (-1,65 kcal/mol). V komplexu s MPER je interakce s CArg683 velice silná a příznivá, zde interaguje arginin zanedbatelně a celková bilance je nevýhodná.

Interakce s Pro100f je zde stabilizující (-3,20 kcal/mol) ale slabší než v MPER (-9,32). Její nejsilnější složkou je CHis60 (-2,32). Ostatní aminokyseliny poskytují slabé interakce, rozložené po velké části interakčního povrchu.

Interakce s AGlu100ije zde zanedbatelně stabilizující.

Z hlediska aminokyselinových záměn prvního klastru (ATPS) je zde pouze zanedbatelná interakce CAla34-APhe100a (-0,0311). Rezidua druhého klastru (GHE) interagují silně s oblastmi jak hydrofobní štěrbinou, tak s počátkem oblasti 1. CGly63-AGlu53 má volnou energii -2,74 kcal/mol; Chis64-AGlu53 -1,53 kcal/mol; CGlu65-APhe100a -5,08. Destabilizující je interakce CGlu65-ATyr99 o volné energii 4,41 kcal/mol. Třetí klastr (VILIL) je v této orientaci inertní.

Jediná interakce MPER-10E8 která je modelem 5 komplexu MLA025-10E8 obdobná je vazba na APhe100A v oblasti 1 protilátky. Existují silné stabilizující interakce se 3/5 reziduí oblasti 1. Druhý klastr aminokyselinových záměn (GHE) se účastní vazby jak s hydrofobní štěrbinou (AGlu53) tak s oblastí 1 (APhe100a stabilizující, ATyr99 destabilizující).

5.2.9.1.3 *MLA092 mod. 0*

S rezidui hydrofobní štěrbinou interaguje MLA092 v predikované orientaci (model 0) zanedbatelně destabilizačně (ATrp33), vůbec (AArg50) a oproti MPER oslabeně stabilizačně (AGlu53 -8,46 kcal/mol oproti MPER, kde je bilance -21,4 kcal/mol). Interakce s AGlu53 je obdobně jako v případě MPER převážně elektrostatická (-8,21 kcal/mol pro model 0 MLA092). Vedená je zejména CTrp37 (-4,84), CGly38 (-3,09) a CAsn39 (-1,44). To poněkud kopíruje vazbu MPER, kde jsou hlavními rezidui CAsn671-CTrp672-CPhe673. Zastoupeny jsou v obou případech blízko sebe ležící Trp a Asn. Phe je zde nahrazen v nejlepší trojici aminokyselinou Gly.

Pro oblast jedna jsou všechny bilance stabilizující s výjimkou APro100f a AGlu100i (mírná destabilizace).

S ATyr99 interaguje MLA025 zhruba o polovinu slaběji (-5,63 vs. -10,5) než v případě MPER. Hlavní je převážně vdW interakce s CAsn39 (-3,52 kcal/mol), oproti MPER interakci čtyřmi sousedními hydrofobními rezidui Leu, Trp, Thr, Trp.

Interakce s APhe100a jsou silnější než v případě MPER (-21,1 vs -10,9 kcal/mol) a spíše elektrostatické. Vede interakce s CLys41 (-11,6), dále jsou významné CAsn43 (-3,90),

CGlu53 (-3,36) a CVal42 (-2,32). Zastoupen zde není výrazně žádný arginin, na rozdíl od vazby s MPER.

Interakce s AGly100d je podstatně slabší než s MPER (-2,05 vs. -17,0), interakce s APro100F i AGlu100i jsou mírně až zanedbatelně destabilizující.

První klastr aminokyselinových záměn (EIMW) se účastní vazby zanedbatelně skrze druhou a třetí aminokyselinu, silně pak skrze čtvrtý tryptofan – interakce CTrp37-AGlu53 s hydrofobní štěrbinou nese volnou energii -4,84 kcal/mol. Druhý klastr (PSW) interaguje pomocí CPro63 s ATrp33 (-0,475) a s AGlu100i (-0,342); skrz CSer64 s ATrp33 (-0,894) a s AGlu100i (-0,122). Z třetího klastru (IVTPL) interagují zanedbatelně všechna rezidua s oblastí 1, nejsilnější je interakce Cval90-ATyr99 (-0,167).

Pro model MLA092 je částečně mimikováno složení interakčních párů s AGlu53 hydrofobní štěrbinou (interakce jsou však oslabeny). V oblasti 1 jsou silné interakce MLA025 s 3/5 reziduí, interakce jsou však oproti MPER slabší. První klastr aminokyselinových záměn interaguje s hydrofobní štěrbinou, druhý interaguje slabě s hydrofobní štěrbinou i oblastí 1. Třetí klastr interaguje slabě až zanedbatelně s oblastí 1.

5.2.9.1.4 MLA158 mod. 2

Celková volná energie interakcí s důležitými rezidui je poměrně slabá – činí -13,3 kcal/mol (oproti MPER -82,5 kcal/mol).

Interakce v tomto modelu komplexu jsou v oblasti hydrofobní štěrbinou slabé (ATrp33 -0,360 kcal/mol) nebo destabilizující (AGlu53 2,44 kcal/mol). S AGlu53 interaguje CLys41 (-3,53) a CAsn43 (-2,69), avšak celkovou vazbu MLA092 na AGlu53 destabilizuje CGln88 (3,55), CGlu50 (3,11) a CTyr90 (2,13).

Interakce s oblastí 1 protilátky mají celkovou volnou energii -15,3 kcal/mol, výrazně slabší než -61,1 pro MPER. Významné jsou interakce s ATyr99 (s CLys41 -2,83; s CGlu53 -1,79) s bilancí -6,89 kcal/mol a s APhe100a (CIle61 -3,01, CTrp36 -1,52), v obou případech zejména van der Waalsovského charakteru. Mírně destabilizující jsou

celkové interakce s APro100f (1,14), stabilizující s AGlu100i (-1,26, zejména s CLys41 -2,73).

První klastr (WMWW) interaguje skrze CMet35-APhe100a (0,136) mírně destabilizačně. CTrp36 interaguje s APhe100a (-1,52) i s ATyr99 (-0,254). Druhý klastr (ITL) interaguje zejména skrze Cile63 s APhe100a (-1,08) a zanedbatelně s ATyr99 (-0,0341). CThr64-APhe100a (0,170) je slabá destabilizující interakce. Třetí klastr (LYYAW) interaguje nepříznivě s hydrofobní štěrbinou – nejsilnější destabilizující interakce jsou CLeu89-AGlu53 (0,310), CTyr90-AGlu53 (2,13), CAla92-AGlu53 (1,21) a CTrp93-AGlu53 (1,61). Nejsilnější stabilizující interakce je CTrp93-ATrp33 (-0,363).

Pro model 2 MLA158 jsou interakce s hydrofobní štěrbinou destabilizující, výrazné stabilizující interakce jsou se 2/5 aminokyselin oblasti 1, slabé s 2/5 aminokyselin a slabě destabilizující s jednou aminokyselinou oblasti 1. První i druhý klastr interaguje příznivě s oblastí 1. Třetí klastr interaguje nepříznivě s hydrofobní štěrbinou, zejména s AGlu53.

5.2.9.1.5 MLA158 mod. 3

Tento model ze všech jako jediný poskytuje výrazné interakce s hydrofobní štěrbinou skrze pozice aminokyselinových záměn a může odpovídat teoretické představě o „smyčkovém vazebném módu“. Vazby na 10E8 protilátku zde tvoří všechny tři regiony aminokyselinových záměn, jak je patrné z vizualizace párových interakcí na **obrázku 31**. (pro instrukce k orientaci v této reprezentaci viz **příloha B 1**) Reprezentace ukazuje, že jsou interakce zejména skrze aminokyselinové záměny na MLA158.

Celková bilance interakcí MLA158-10E8 fr. je v tomto modelu -12,1 kcal/mol (oproti MPER -82,45 kcal/mol). Silné jsou interakce s hydrofobní štěrbinou (-10,8) přibližně stejně elektrostatického (-5,56) a van der Waalsovského charakteru (-5,28). Nejsilněji interaguje MLA158 s ATrp33 (-6,04), zejména pomocí dvou sousedních reziduí CTrp37 (-4,08, převážně vdW) a CTrp36 (-1,09, převážně elstat.), což koresponduje s tryptofanem CTrp672 z MPER interagujícím van der Waalsovsky.

Interakce MLA158 s AArg50 je, stejně jako u MPER destabilizující (0,0465), oproti MPER (3,25) však zanedbatelně. Významná je i interakce s AGlu53 (-4,85), pocházející

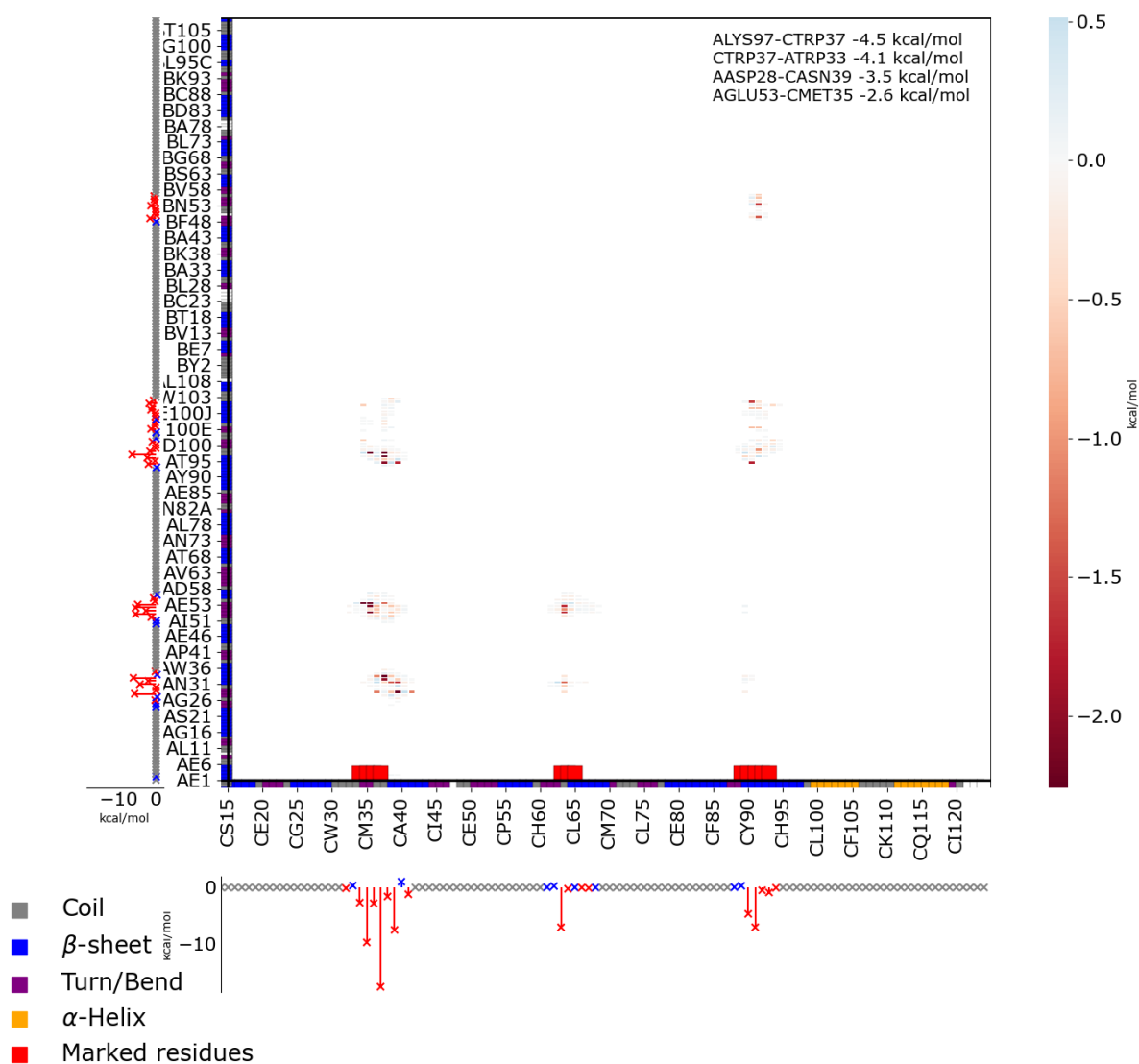
převážně od reziduí CMet35 (-2,64), CTrp34 (-2,26) a CTrp37 (-1,32). Oproti MPER je vazba na toto reziduum zhruba 4× slabší, zhruba 2× silnější je však vazba na ATrp33.

Interakce s oblastí 1 jsou slabé, interakce s AGly100d je zanedbatelně destabilizující a neexistuje žádná energeticky význačná interakce.

První klastr (WMWW) interaguje převážně skrze interakce CTrp34-AGlu53 (-2,26), CTro34-APhe100a (-0,538), CMet35-AGlu53 (-2,64), CMet35-ATyr99 (-0,160), CTrp36-ATrp33 (-1,09) a nejsilněji skrze CTrp37-ATrp33 (-4,08). Dále existuje stabilizující interakce CTrp37-AGlu53 (-1,32). Klastr 2 (ITL) interaguje slabě s hydrofobní štěrbinou (bilance je mírně destabilizující). Klastr 3 (LYYAW) interaguje s důležitými rezidui 10E8 zanedbatelně. Nejvýznačnější interakcí s hydrofobní štěrbinou je CLeu89-ATrp33 (-0,123). S oblastí 1 je zde řada zanedbatelných interakcí stabilizujícího i destabilizujícího charakteru, z nichž nejsilnější stabilizující je CAla92-ATyr99 (-0,350), nejsilnější destabilizace je pak způsobená párem CTrp33-ATyr99 (0,285).

MLA158-10E8fr. model 3 obsahuje předpokládaný smyčkový vazebný mód, kde je vazba MLA158 na protilátku vedena všemi klastry zaměněných aminokyselin. Párové interakce s hydrofobní štěrbinou jsou dobré, MLA158 mod. 3 společně s mod. 8 jako jediné mají interakce s AArg50, byť zanedbatelně destabilizující. Interakce s oblastí 1 protilátky jsou nevýrazné a nedomnuje jim výrazně žádná aminokyselina. Klastr 1 interaguje silně s hydrofobní štěrbinou. Klastry 2 a 3 interagují s 10E8 jen slabě.

10E8-MLA158 mod. 3 total_IEM



Obrázek 31 – Reprezentace vazby MLA158 na 10E8fr. z MD trajektorie modelu 3 MLA0158-10E8.

5.2.9.1.6 MLA158 mod. 8

Celková energie párových interakcí s důležitými rezidui protilátky činí -33,7 kcal/mol. Tento model následuje β -vazebný mód, s hydrofobní štěrbinou interaguje zanedbatelně (-0,0485 kcal/mol) – s ATrp33 a AArg50 jsou bilance interakcí mírně destabilizující (0,0283; 0,240 kcal/mol), interakce s AGlu53 je mírně stabilizující (-0,517 kcal/mol). S AGlu53 existuje velmi silná interakce AGlu53-CArg28 (-25,79) elektrostatické povahy, vyvážená silnou destabilizující elektrostatickou interakcí AGlu53-CGlu20 (20,7 kcal/mol). Destabilizující je dále elektrostatická interakce AGlu53-CGlu69 (7,54), stabilizující interakce existují s CGln32 (-1,30), CLeu65 (-1,86) a Cile67 (-2,19). Silné jsou však interakce s oblastí 1 protilátky (-33,7 kcal/mol), zhruba stejně elektrostatického a vdw charakteru.

S ATyr99 poskytuje MLA158 celkovou volnou energii interakcí -14,9 kcal/mol, způsobenou zejména CHis60 (-8,32), CGlu69 (-3,21) a CGln26 (-2,89). Pro porovnání jsou interakce tohoto rezidua s MPER převážně vdw, vedené rezidui Leu, Trp a Thr.

Nejsilnější jsou celkově interakce s APhe100a (-20,2 kcal/mol, spíše van der Waalsovského charakteru), hlavně s CLys58 (-7,47), CGlu73 (-7,53), CPhe71 (-3,06) a CGln26 (-2,89). Zbytek interakcí je buď slabý/zanedbatelný nebo destabilizující (AGly100d 0,0523; APro100f -0,964; AGlu100i 2,29)

Klastr 1 (WMWW) zde interaguje jen slabě s AGlu53 hydrofobní štěrbinou (skrze CTrp34 a CMet35: -0,110; -0,173). Role aminokyselin klastru 2 (ITL) je zanedbatelná s výjimkou CLeu65, který interaguje s hydrofobní štěrbinou skrze CLeu65-ATrp33 (-0,361) a CLeu65-AGlu53 (-1,86). Klastr 3 (LYYAW) se vazby s důležitými rezidui 10E8 vůbec neúčastní.

Interakce MLA158 s hydrofobní štěrbinou jsou v tomto modelu celkově zanedbatelné (byť existují velmi silné interakce s AGlu53, které jsou však vyváženy silnými destabilizujícími interakcemi). Existují silné interakce s prvními dvěma rezidui oblasti 1 (ATyr99, APhe100a) odlišného charakteru než u MPER. Role klastru 1 je z energetického hlediska zanedbatelná a směřovaná na hydrofobní štěrbinu. Role CLeu65 klastru 2 spočívá ve stabilizující vazbě na AGlu53 hydrofobní štěrbinou. Klastr 3 se vazeb vůbec neúčastní.

5.2.9.1.7 *MyoWT*

Interakce kontrolních MyoWT proteinů s důležitými rezidui 10E8 jsou pro model 6 -88,3 kcal/mol. Celkově jsou tyto interakce tvořeny primárně interakcí AGlu100i-CLys41 (-83,5 kcal/mol) – stejnou interakcí s neobvykle vysokou predikovanou volnou energií, jako u vyřešené struktury MPER. Byť jsou ve fyziologických podmínkách obě aminokyselinová rezidua nabitá opačnými náboji a dle simulací mají velmi blízkou interakci (v průběhu trajektorie směřují k sobě), názorem autora je, že je energie těchto interakcí nadhodnocená. Při zanedbání této interakce lze říct, že MyoWT model 6 má ve struktuře nulovou interakci s hydrofobní štěrbinou mezi MyoWT a 10E8fr., slabé stabilizující interakce s AGly100d (-0,392) a APro100f (-1,62) a slabě destabilizující interakci s ATyr99 (1,25 kcal/mol). Jediné významně stabilizující interakce existují s APhe100a (-5,78), zejména skrze CArg63 (-6,14) – podobná interakce existuje i ve vyřešené struktuře MPER.

MyoWT v modelu 8 interaguje s hydrofobní štěrbinou spíše neochotně (10,6 kcal/mol), zejména díky interakci s AGlu53 (10,6) skrze CGlu20 (10,9). S oblastí 1 protilátky existují ze všech modelů nejsilnější interakce s ATyr99 (-23,2) převážně elektrostatického charakteru s CGlu69 (-17,46). Stabilizující jsou i interakce s APhe100a (-9,88), zejména s CPhe71 (-3,83), CGlu69 (-1,94) a řadou dalších aminokyselin. Destabilizující jsou všechny ostatní bilance interakcí s důležitými rezidui – AGly100d (2,98), APro100f (3,36) a AGlu100i (8,36).

Klastr 1 (EKLS) modelu 6 má pouze zanedbatelné interakce skrze CLeu36 do oblasti 1 a slabé destabilizující interakce s oblastí 1 skrze CSer37. Klastr 2 (RNT) má jedinou významnou silnou stabilizující interakci CArg63-APhe100a (-6,14). Klastr 3 (DGKAT) má interakce zanedbatelné a jen skrze CAsp89 s bilancí přibližně rovnou nule.

Díky odlišnému místu vazby má model 8 nulové interakce skrze klastr 1, zanedbatelné (stabilizující) interakce skrze klastr 2 a žádné interakce skrze klastr 3.

Ani jeden z modelů vazby MyoWT-10E8 nemá příznivé interakce s hydrofobní štěrbinou (MyoWT mod. 6 žádné; MyoWT mod. 8 destabilizující). Existují stabilizující interakce s oblastí 1 (MyoWT mod. 6 4/5 reziduí; MyoWT mod. 8 2/5) které jsou však slabší než u ostatních zkoumaných modelů proteinů. Z hlediska klastrů aminokyselinových záměn jsou všechny interakce obou modelů nulové nebo zanedbatelné s jedinou výjimkou silné interakce CArg63-APhe100a u modelu 6.

5.2.10 Souhrn analýz intermolekulových interakcí

V předchozích kapitolách byly zkoumané modely struktur komplexů MLA/MyoWT/MPER zkoumány z hlediska svých párových interakcí známými důležitými rezidui protilátky 10E8. Analyzovány byly i z hlediska klastrů aminokyselinových záměn.

Kontrola vybraného modelu komplexu s MPER (mod. 6) byla porovnána s vyřešenou strukturou. Nejprve byly identifikovány klíčové interakční páry s protilátkou na základě svých interakčních volných energií. Zde došlo ke shodě nejdůležitějších reziduí MPER identifikovaných těmito kritérii s rezidui označenými jako klíčovými autory struktury. Pro dokování a metodou MM/PBSA vybraný model 6 stejného komplexu byly identifikovány velmi podobné interakce. Spojení tohoto zjištění, přijatelných hodnot RMSD (byť je první helix modelu oproti vyř. str. vyosen směrem od proteinu) byl dokováním získaný model označen jako kvalitní – tento způsob kontroly ukázal, že je (minimálně v tomto případě) výsledek dokování pomocí ClusPro2 v souladu s vyřešenou krystalovou strukturou.

Dále byly pro 6 tzv. β -vazebných módů (alternativ k hypotetickým smyčkovým módům) identifikovány mimikující interakce – takové, které by kopírovali interakce MPERu s protilátkou. Totéž bylo provedeno pro 2 smyčkové módy modely MLA158 (módy s účastí 3 regionů aminokyselinových záměn) – mód 2 s vazbou pod úhlem, mód 3 s vazbou v ose.

Bylo zjištěno, že jsou zkoumané β -vazebné módy schopné příznivě interagovat s oblastí 1 protilátky, zejména s rezidui ATyr99 a APhe100a. Výrazné interakce naopak chyběly s AGly100d, APro100f a AGlu100i – interakce, které jsou s MPER přítomny. Celkové síly párových interakcí poskytnutých β -vazebnými módy byly -12,11 až -36,21 kcal/mol. Interakce skrze klastry aminokyselinových záměn nebyly vůbec pozorovány pro model MLA024 a modely MyoWT (s jedinou výjimkou silné interakce CArg63-APhe100a pro MyoWT model 6). Pozorovány byly naopak u všech ostatních vybraných modelů, zejména skrze první a druhé klastry směřované na hydrofobní štěrbinu 10E8. Jedinou

výjimkou byl model 2 vazby MLA158 jehož stabilizující interakce směřovaly primárně do oblasti 1 protilátky.

Díky náznaku mimikujících schopností skrze klastry aminokyselinových záměn se nabízí diskuze, zda-li nejsou zde popsané „ β -vazebné módy“ ve skutečnosti těmi, které popisují reálnou interakci MLA proteinů s protilátkou (ať už *in vitro*, či v biologickém kontextu). Z fyzikálně-chemického hlediska a z kontextu celého povrchu interakce jsou stabilní (s příznivou vazebnou ΔG) a interagují (na rozdíl od dosud hypotetických smyčkových vazebných módů) pomocí velkých interakčních povrchů a řady párových interakcí (porovnání vazebných módů poskytují výsledky US mezi modely 8 a 3 MLA158). Z hlediska možné biologické aktivity ve všech modelech (s výjimou MLA24 a MyoWT) existují interakce 10E8fr. s klastry aminokyselinových záměn – předpokládaným mechanismem účinku *in vivo*. Dle výsledků zde blíže nepopsaných *in vivo* experimentů je zřejmé, že pro mimimování závisí opravdu na aminokyselinových záměnách (MyoWT v těchto experimentech nevyvíjelo žádnou aktivitu).⁵¹

Pro smyčkové vazebné módy MLA158 byly v případě modelu 3 (interakce v ose/smyčkový mód) byly jako pro jediný model nalezeny v bilanci stabilizující interakce s rezidui hydrofobní štěrbinou protilátky, interakce s oblastí 1 byly naopak slabě stabilizující.

V případě modelu 2 vazby MLA158 jsou naopak interakce s hydrofobní štěrbinou zanedbatelné, silnější jsou interakce s oblastí 1 a celý MLA158 protein je „vyosen“ oproti ose protilátky. Identifikovat „silnější“ z těchto dvou orientací MLA158 na základě zde popsáných výsledků nelze zcela s jistotou, dle atomistických US simulací trajektorie odsedání proteinu od protilátky se však zdá, že model 3 ($\Delta PMF = 59,61$ kJ/mol) bude existovat s vyšší pravděpodobností, než model 2 ($\Delta PMF = 53,72$ kJ/mol). Oproti tomu MLA158 v modelu 8 (β) má nejsilnější US stanovenou vazebnou volnou energii ($\Delta PMF = 61,8$ kJ/mol) a měla by tedy z tohoto hlediska být ze zde tří popsáných orientací nejpravděpodobnější a nejsilnější.

Analýzy ukázaly i schopnost MyoWT vázat některá rezidua protilátky, ani jeden z modelů však nevykazoval vazebné schopnosti k hydrofobní štěrbině. Model 6 vykazoval silnou interakci s APhe100a a s AGlu100i (zde je však celková síla stále otázkou). Model 8 pak silnou interakci s ATyr99 (silnější než u jakéhokoliv jiného analyzovaného modelu) a AGly100d, avšak destabilizující interakce se zbytkem aminokyselin v oblasti 1. Oba modely jsou dle přibližných CG US simulací i MM/PBSA schopny stabilně a silně vázat protilátku (MyoWT mod. 6 Δ PMF = 86,18 kJ/mol avšak s problematickou AGlu100i-CLys41 vazbou; MyoWT mod. 8 Δ PMF = 84,94 kJ/mol). Schopnost MyoWT vázat protilátku 10E8 byla prokázána i v experimentální části této práce.

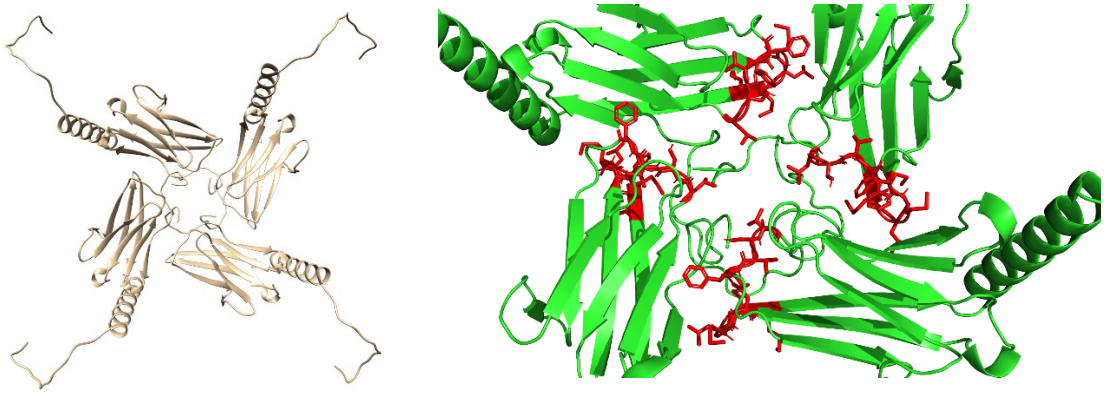
5.2.11 Tetramer MLA proteinů

Na základě experimentálního objevu tetramerních forem MLA proteinů a zdánlivému nesouladu pořadí eluce Hsp70 a MLA-tetramer frakcí v SEC bylo provedeno přibližné modelování tetramerní struktury (včetně tagů) pomocí prvotního homologního modelu MLA024 a webového serveru GalaxyHomomer. Struktura byla simulována pomocí 20 ns MD simulace, během které se jevila jako stabilní (výsledky nejsou v této práci uvedeny). Pomocí aplikace HYDROPRO byl vypočten její translační difuzní koeficient, který činil $D_{t,MLA024-tetramer} = 4,80 \text{ cm}^2 \cdot \text{s}^{-1}$ pro tento 64 kDa protein. Obdobně byl vypočten i $D_{t,Hsp70} = 5,49 \cdot 10^{-7} \text{ cm}^2 \cdot \text{s}^{-1}$ pro strukturu bakteriálního proteinu Hsp70 získanou z PDB databáze pod kódem 2KHO.⁸⁷ Tento výsledek byl ověřen stejným výpočtem na základě struktury predikované softwarem AlphaFold z příslušné strukturní databáze.⁹¹ Hodnota pro predikovanou strukturu činila $D_{t,Hsp70,II} = 5,45 \cdot 10^{-7} \text{ cm}^2 \cdot \text{s}^{-1}$.

Vzhledem k závislosti translačního difuzního koeficientu k hydrodynamickému poloměru vyjádřeném Stokes-Einsteinovou rovnicí, ze které plyne že (R_H označuje Stokesův poloměr (hydrodynamický poloměr)) je hydrodynamický poloměr Hsp70 menší, než MLA024-tetrameru. Toto vysvětluje, proč při SEC purifikaci putuje tetramer jako první, ačkoliv je jeho molekulová hmotnost nižší.

$$R_h \propto \frac{1}{D_t}$$

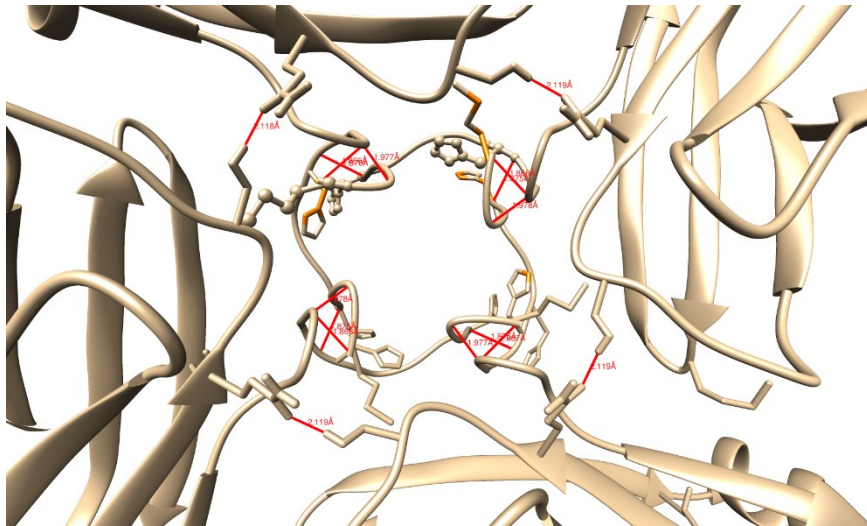
Podoba struktury je zobrazena **obrázku 32**. Možné vodíkové vazby mezi monomery komplexu jsou viditelné na **obrázku 33**.



Obrázek 32 – Model struktury tetrameru MLA024.

Vlevo) Celkový pohled zeshora

Vpravo) pohled na aminokyselinové záměny



Obrázek 33 – Možné vodíkové vazby mezi monomery MLA024-tetrameru.

Struktura je relativně planárního charakteru (jak naznačuje zvýšený Stokesův poloměr) a může být stabilizována mj. četnými vodíkovými vazbami mezi N-terminálními tagy a krátkou částí samotného proteinu. Jednoznačné je, že byla nějaká forma tetramerní struktury pozorována v *in vitro* experimentech, jak dokazují HPLC chromatogramy a DLS experimenty.

Omezeným experimentálním ověřením formace tetrameru skrze N-terminální části proteinů jsou nepublikované výsledky HPLC a DLS analýz na konstruktech His-TEV-MLB-V5 proteinů knihovny MLB založených též na myomedinech. V HPLC chromatogramech těchto proteinů je patrná absence píků odpovídajících tetramerním

strukturám. Dva vybrané chromatogramy jsou uvedeny v **přílohách B 24 a B 25** a náleží Bc. Tereze Zosinčukové (Laboratoř inženýrství vazebných proteinů, BTÚ AV ČR). Použité konstrukty obsahovaly kromě His-tagu zároveň i místo pro štěpení TEV-proteasou, které mění geometrii N-terminální části proteinu. Pokud by došlo k narušení geometrie potřebné k vytvoření tetrameru tak, jak je zde prezentován, vysvětlilo by to jeho absenci v HPLC chromatogramech a následných DLS (výsledky nepublikovány).

5.3 Souhrn a návrhy na vylepšení vazebných proteinů

Autor této práce je názoru, že jsou zde popsány β -vazebné módy pravděpodobně skutečnými popisy interakcí mezi MLA proteiny a 10E8 protilátkou. Konkrétní popsané struktury nemusí odpovídat realitě (bylo by třeba je ověřit experimentálně), avšak je spolehlivě prokázáno, že mohou vazebné módy tohoto typu mít vysokou afinitu, nést vhodné interakce s důležitými rezidui protilátky a četné další stabilizující interakce. Podobné orientace MLA proteinu skrze tyto módy pravděpodobně tvoří pevnější vazby se zachováním chtěné biologické aktivity. Zde popsány vazebné módy byly vybrány pro každou variantu z 10 módů získaných dokováním, následovala hrubá selekce metodou MM/PBSA a přesnější selekce pomocí CG US. Zejména MM/PBSA metoda je aproximativní, byť vhodná na podobné selekce ligandů k jednomu cíli, a nemohla být z hlediska parametrů (zejména permitivity vazebného povrchu) kalibrována s využitím experimentálních dat. Pokud byl obsažen opravdový vazebný mód v prvních deseti modelech, mohl se zejména v tomto kroku selekce ztratit. Nutno je říci, že zde popsaná selekce na základě vazebných Gibbsových energií nemusí v principu nutně vybrat strukturu, která existuje v nativních podmínkách.

Dalším nepřímým důkazem pro vazbu větším povrchem byla skutečnost, že výtěžek selekce Myomedinové knihovny ve vazbě na 10E8 IgG-1 byl až překvapivě vysoký oproti jiným terčům. Byla porovnána data úspěšnosti selekce téže Myomedinové knihovny pomocí ribozomálního displeje proti třem různým terčům IgG-1, konkrétně 10E8, PGT121 a PGT126. V případě 10E8 bylo z celkového počtu 300 klonů pomocí ELISA selektováno 30 variant, které specificky rozpoznávají neutralizující protilátku, pro další charakterizaci bylo vybráno 8 nejlepších kandidátů. Zatímco u PGT121 a PGT126 byly nalezeny jen 3 u každé protilátky, které se ukázaly jako specifické pro danou protilátku. Z vazebných křivek je rovněž patrné, že žádná z variant vážící PGT121 i PGT126 nedosahuje K_d nejlepších variant u 10E8, afinita v podobě K_d byla odhadována na základě inflexních bodů z ELISA (Kuchař, ústní sdělení).

V rámci *in silico* výpočetní části byla identifikována dvě další rezidua ligandu CThr676 a CAsn671 MPER (na základě vyřešené struktury 4G6F komplexu), která by mohla být významná pro interakci protilátky 10E8 s gp41 viru HIV-1.

Pokud jsou *in silico* části popsané vazebné módy opravdu těmi, které existují (byť může ve skutečnosti existovat rovnováha mezi různými konformacemi) bylo by na místě otestovat vazebné proteiny s úpravami v aminokyselinových klastrech. Například odstraněním třetího klastru záměnných aminokyselin z MLA158 by mohlo být teoreticky dosaženo velmi podobné vazby, ale stabilita proteinu v roztoku by se mohla zvýšit, díky odstranění (namutování zpět na WT sekvenci) výrazně hydrofobní části proteinu exponované do rozpouštědla. Odlišným postupem k možnému vylepšení by bylo zkusit zlepšit interakce MLA proteinů s hydrofobní šterbinou protilátky 10E8 pomocí cílených zásahů do sekvence.

Pro snížení vazby MyoWT na 10E8 (byť jen jako potvrzení teoretického vazebného módu) by bylo možno odstranit CLys41, napěstovat protein a provést experiment ke stanovení vazebné afinity na 10E8.

Je nutné zmínit i možná vylepšení, která mohou být vnesena do této práce v případě jejího opakování a upřesnění. Z hlediska experimentálního by bylo vhodné provést měření afinit přesnějšími metodami v roztoku, vhodná by byla metoda MST. Zároveň by bylo přínosné zjistit skutečnou podobu vazby MLA proteinů na 10E8 protilátku, vzhledem k problémům s krystalografickým stanovením se nabízí metoda SAXS spojená s porovnáním experimentální křivky s teoretickými strukturami prezentovanými v této práci. Z hlediska *in silico* části by pro validaci výsledků bylo vhodné provést delší MD simulace, jelikož byl díky kratším trajektoriím v této práci omezen počet analyzovatelných snímků. Zároveň by bylo vhodné simulace duplikovat a provést statistickou analýzu získaných výsledků.

6 Závěr

Práce měla dvě distinktní části. V rámci *in vitro* experimentální části byly z knihovny plasmidů nesoucích sekvence potenciálních vazebných proteinů MLA vyselektovány varianty s nejpříznivějšími poměry vazby na 10E8 protilátku a isotypovou kontrolní protilátku IgG λ . Tyto vybrané proteiny byly analyzovány z hlediska vazebných afinit na obě zmíněné protilátky. U nejlepších z analyzovaných proteinů byl molekulárně biologickými metodami vyměněn C-terminální Avi-tagu za V5-tag, tím byl vytvořen konstrukt His-MLA-V5. Proteiny byly exprimovány v kulturách *E. coli*, purifikovány afinitní chromatografií a analyzovány z hlediska vazebných afinit. Proteiny s nejlepším poměrem vazby IgG λ /10E8 a nejlepší afinitou k 10E8 byly odeslány na *in vivo* vakcinační experimenty.

Na základě *in vivo* výsledků bylo vybráno několik málo proteinů pro další analýzy. Na těchto proteinech MLA024, 025, 092 a 158 byly provedeny purifikace s využitím gelové permeační chromatografie. U všech variant, včetně MyoWT, byly identifikovány tři charakteristické frakce produkovaných proteinů. Jedna z těchto frakcí odpovídá očekávanému monomernímu produktu, další frakce sestává z Hsp70-MLA komplexu kopurifikovaného na Ni-NTA a třetí frakce je pravděpodobně oligomerním stavem myomedinu. Na některých frakcích variant byly provedeny základní biofyzikálně-chemické analýzy, které zjistily teploty tání proteinů, jejich hydrodynamické poloměry a sekundární strukturu.

In silico část této práce se zabývala stejnými proteiny z knihovny MLA. Byly vybudovány homologní modely, které byly vylepšovány pomocí unikátního protokolu třech po sobě jdoucích krátkých MD simulací. Tyto modely byly dokovány společně s vyřešenou strukturou 10E8 protilátky, vybrané modely byly dále selektovány MD simulací a MM/PBSA analýzou. U takto vybraných modelů byla provedena analýza volných energií vazby pomocí coarse-grained Umbrella Sampling simulací, u modelů vazeb proteinu MLA158 byly provedeny i simulace s atomistickým rozlišením.

Vybrané modely vazeb MLA-10E8 byly analyzovány tak, že byly snímky z jejich MD trajektorií využity k výpočtu volných energií párových interakcí aminokyselin. Tyto

energie byly vizualizovány v podobě dávající kontext energií k sekundárním strukturním elementům, k řetězcům proteinů a k celkovým příspěvkům jednotlivých částí obou vazebných partnerů. Na základě matic volných energií párových interakcí proběhlo srovnání vazby nativního komplexu 10E8-MPER, jeho modelů vybraných identickou cestou jak u MLA proteinů, a vybraných modelů vazeb 10E8-MLA. Toto srovnání se zaměřovalo na interakce kopírující nativní vazby 10E8-MPER, na interakci důležitých reziduí dvou regionů protilátky s MLA proteiny.

Budoucnost výzkumu spočívá nyní převážně v detailní charakterizaci vybraných vazebných proteinů a v jejich vylepšování z hlediska stability a vazebných vlastností. Zde vybrané a popsané varianty z knihovny MLA byly použity pro *in vivo* experimenty, kde prokázaly schopnost sér vakcinovaných myší neutralizovat až 54 % HIV-1 pseudovirů z reprezentativního panelu.⁵¹

7 Seznam literatury

1. de Silva, T. I., Cotten, M. & Rowland-Jones, S. L. HIV-2: the forgotten AIDS virus. *Trends Microbiol.* **16**, 588–595 (2008).
2. WHO | Data and statistics. *WHO* <http://www.who.int/hiv/data/en/>. (přístup 27.4.2020)
3. Maartens, G., Celum, C. & Lewin, S. R. HIV infection: epidemiology, pathogenesis, treatment, and prevention. *Lancet* **384**, 258–271 (2014).
4. Korber, B. *et al.* Timing the ancestor of the HIV-1 pandemic strains. *Science* **288**, 1789–1796 (2000).
5. Faria, N. R. *et al.* HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science* **346**, 56–61 (2014).
6. Barré-Sinoussi, F. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868–871 (1983).
7. Vella, S., Schwartländer, B., Sow, S. P., Eholie, S. P. & Murphy, R. L. The history of antiretroviral therapy and of its implementation in resource-limited areas of the world. *AIDS* **26**, 1231–1241 (2012).
8. Furman, P. A. *et al.* Phosphorylation of 3'-azido-3'-deoxythymidine and selective interaction of the 5'-triphosphate with human immunodeficiency virus reverse transcriptase. *Proc Natl Acad Sci U S A* **83**, 8333–8337 (1986).
9. Vella, S. *et al.* Survival of Zidovudine-Treated Patients With AIDS Compared With That of Contemporary Untreated Patients. *JAMA* **267**, 1232–1236 (1992).
10. Larder, B. A., Darby, G. & Richman, D. D. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. *Science* **243**, 1731–1734 (1989).
11. Skowron, G. *et al.* Alternating and intermittent regimens of zidovudine and dideoxycytidine in patients with AIDS or AIDS-related complex. *Ann. Intern. Med.* **118**, 321–330 (1993).
12. Hammer, S. M. *et al.* A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. AIDS Clinical Trials Group Study 175 Study Team. *N. Engl. J. Med.* **335**, 1081–1090 (1996).
13. Montaner, J. S. *et al.* A randomized, double-blind trial comparing combinations of nevirapine, didanosine, and zidovudine for HIV-infected patients: the INCAS Trial. Italy, The Netherlands, Canada and Australia Study. *JAMA* **279**, 930–937 (1998).
14. Cameron, W. *et al.* Ritonavir and saquinavir combination therapy for the treatment of HIV infection. *AIDS* **13**, 213–224 (1999).
15. Gulick, R. M. *et al.* 3-year suppression of HIV viremia with indinavir, zidovudine, and lamivudine. *Ann. Intern. Med.* **133**, 35–39 (2000).
16. Palella, F. J. *et al.* Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N. Engl. J. Med.* **338**, 853–860 (1998).
17. van Sighem, A. I. *et al.* Life expectancy of recently diagnosed asymptomatic HIV-infected patients approaches that of uninfected individuals. *AIDS* **24**, 1527–1535 (2010).
18. Rosypal, S. *Úvod do molekulární biologie. Třetí díl.* (2002).
19. Human Immunodeficiency Virus (HIV). *Transfus Med Hemother* **43**, 203–222 (2016).
20. PDB101: Learn: Flyers, Posters & Other Resources: The Structural Biology of HIV (Poster). *RCSB: PDB-101* <https://pdb101.rcsb.org/learn/flyers-posters-and-other-resources/poster/the-structural-biology-of-hiv-poster>.
21. Landmarks of the HIV genome. <https://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html>.
22. Levin, J. G., Mitra, M., Mascarenhas, A. & Musier-Forsyth, K. Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. *RNA Biol* **7**, 754–774 (2010).
23. Donahue, J. P., Vetter, M. L., Mukhtar, N. A. & D'Aquila, R. T. The HIV-1 Vif PPLP Motif is Necessary for Human APOBEC3G Binding and Degradation. *Virology* **377**, 49–53 (2008).
24. Smyth, R. P., Davenport, M. P. & Mak, J. The origin of genetic diversity in HIV-1. *Virus Res.* **169**, 415–429 (2012).
25. Cuevas, J. M., Geller, R., Garijo, R., López-Aldeguer, J. & Sanjuán, R. Extremely High Mutation Rate of HIV-1 In Vivo. *PLoS Biology* **13**, e1002251 (2015).
26. Zhang, H. *et al.* The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**, 94–98 (2003).

27. Ji, J. P. & Loeb, L. A. Fidelity of HIV-1 reverse transcriptase copying RNA in vitro. *Biochemistry* **31**, 954–958 (1992).
28. Hu, W. S. & Temin, H. M. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc Natl Acad Sci U S A* **87**, 1556–1560 (1990).
29. Li, G. *et al.* An integrated map of HIV genome-wide variation from a population perspective. *Retrovirology* **12**, 18 (2015).
30. Li, G. *et al.* Functional conservation of HIV-1 Gag: implications for rational drug design. *Retrovirology* **10**, 126 (2013).
31. Bonomelli, C. *et al.* The Glycan Shield of HIV Is Predominantly Oligomannose Independently of Production System or Viral Clade. *PLoS One* **6**, e23521 (2011).
32. Lasky, L. A. *et al.* Neutralization of the AIDS retrovirus by antibodies to a recombinant envelope glycoprotein. *Science* **233**, 209–212 (1986).
33. Korber, B. *et al.* Evolutionary and immunological implications of contemporary HIV-1 variation. *Br Med Bull* **58**, 19–42 (2001).
34. Doores, K. J. The HIV glycan shield as a target for broadly neutralizing antibodies. *FEBS J* **282**, 4679–4691 (2015).
35. Pritchard, L. K. *et al.* Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies. *Nat Commun* **6**, 7479 (2015).
36. Crispin, M. & Doores, K. J. Targeting host-derived glycans on enveloped viruses for antibody-based vaccine design. *Current Opinion in Virology* **11**, 63–69 (2015).
37. Zhu, P. *et al.* Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* **441**, 847–852 (2006).
38. Zanetti, G., Briggs, J. A. G., Grünewald, K., Sattentau, Q. J. & Fuller, S. D. Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ. *PLoS Pathog* **2**, e83 (2006).
39. Burton, D. R. & Hangartner, L. Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design. *Annu Rev Immunol* **34**, 635–659 (2016).
40. Do Kwon, Y. *et al.* Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat Struct Mol Biol* **22**, 522–531 (2015).
41. Broad neutralization coverage of HIV by multiple highly potent antibodies | Nature. <https://www.nature.com/articles/nature10373/>.
42. Tomaras, G. D. *et al.* Initial B-Cell Responses to Transmitted Human Immunodeficiency Virus Type 1: Virion-Binding Immunoglobulin M (IgM) and IgG Antibodies Followed by Plasma Anti-gp41 Antibodies with Ineffective Control of Initial Viremia. *J Virol* **82**, 12449–12463 (2008).
43. Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).
44. Richman, D. D., Wrinn, T., Little, S. J. & Petropoulos, C. J. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc Natl Acad Sci U S A* **100**, 4144–4149 (2003).
45. Hraber, P. *et al.* Prevalence of broadly neutralizing antibody responses during chronic HIV-1 infection. *AIDS* **28**, 163–169 (2014).
46. Moore, P. L., Williamson, C. & Morris, L. Virological features associated with the development of broadly neutralizing antibodies to HIV-1. *Trends Microbiol* **23**, 204–211 (2015).
47. Simek, M. D. *et al.* Human Immunodeficiency Virus Type 1 Elite Neutralizers: Individuals with Broad and Potent Neutralizing Activity Identified by Using a High-Throughput Neutralization Assay together with an Analytical Selection Algorithm. *J Virol* **83**, 7337–7348 (2009).
48. Kwong, P. D., Mascola, J. R. & Nabel, G. J. Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nat Rev Immunol* **13**, 693–701 (2013).
49. Frey, G. *et al.* A fusion-intermediate state of HIV-1 gp41 targeted by broadly neutralizing antibodies. *Proc Natl Acad Sci U S A* **105**, 3739–3744 (2008).
50. Chakrabarti, B. K. *et al.* Direct Antibody Access to the HIV-1 Membrane-Proximal External Region Positively Correlates with Neutralization Sensitivity ν . *J Virol* **85**, 8217–8226 (2011).
51. Kuchař, M. *et al.* Myomedin scaffold variants targeted to 10E8 HIV-1 broadly neutralizing antibody mimic gp41 epitope and elicit HIV-1 virus-neutralizing sera in mice. *Virulence* **12**, 1271–1287 (2021).
52. Gebauer, M. & Skerra, A. Engineered Protein Scaffolds as Next-Generation Therapeutics. *Annual Review of Pharmacology and Toxicology* **60**, 391–415 (2020).
53. Kosztyu, P. *et al.* Proteins mimicking epitope of HIV-1 virus neutralizing antibody induce virus-neutralizing sera in mice. *EBioMedicine* **47**, 247–256 (2019).

54. Besmer, E., Market, E. & Papavasiliou, F. N. The Transcription Elongation Complex Directs Activation-Induced Cytidine Deaminase-Mediated DNA Deamination. *Mol Cell Biol* **26**, 4378–4385 (2006).
55. QtiPlot. <https://qtiplot.com/index.html>.
56. ExpASy - ProtParam tool. <https://web.expasy.org/protparam/>.
57. UniProt: the universal protein knowledgebase in 2021 | Nucleic Acids Research | Oxford Academic. <https://academic.oup.com/nar/article/49/D1/D480/6006196>.
58. QIAprep Spin Miniprep Kit. <https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/plasmid-dna/qiaprep-spin-miniprep-kit/>.
59. QIAquick Gel Extraction Kit. <https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/dna-clean-up/qiaquick-gel-extraction-kit/>.
60. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779–815 (1993).
61. Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein science* **27**, 293–315 (2017).
62. Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011).
63. *The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.* <https://www.pymol.org>
64. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612 (2004).
65. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33–38, 27–28 (1996).
66. D.A. Case, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, *et al.* *AMBER 2020*. (University of California, 2020).
67. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43–56 (1995).
68. Spoel, D. V. D. *et al.* GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry* **26**, 1701–1718 (2005).
69. Huang, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods* **14**, 71–73 (2017).
70. Hornak, V. *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725 (2006).
71. Machado, M. R. *et al.* The SIRAH 2.0 Force Field: Altius, Fortius, Citius. *J. Chem. Theory Comput.* **15**, 2719–2733 (2019).
72. GitHub - tildeslu/g_mmpbsa: MM-PBSA method for GROMACS. For full description, please visit homepage: [GitHub](https://github.com/tildeslu/g_mmpbsa) https://github.com/tildeslu/g_mmpbsa.
73. Mario Sergio Valdés Tresanco, Mario E. Valdés-Tresanco, Pedro A. Valiente & Ernesto Moreno Frias. *gmx_MMPBSA*. (Zenodo, 2021). doi:10.5281/zenodo.4732765.
74. Serçinoglu, O. & Ozbek, P. gRINN: a tool for calculation of residue interaction energies and protein energy network analysis of molecular dynamics simulations. *Nucleic Acids Res* **46**, W554–W562 (2018).
75. Hub, J. S., de Groot, B. L. & van der Spoel, D. g_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *J. Chem. Theory Comput.* **6**, 3713–3720 (2010).
76. Miller, B. R. *et al.* MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J Chem Theory Comput* **8**, 3314–3321 (2012).
77. Kozakov, D. *et al.* The ClusPro web server for protein-protein docking. *Nat Protoc* **12**, 255–278 (2017).
78. S, V. *et al.* New additions to the ClusPro server motivated by CAPRI. *Proteins* **85**, 435–444 (2017).
79. Kozakov, D. *et al.* How good is automated protein docking? *Proteins: Structure, Function, and Bioinformatics* **81**, 2159–2166 (2013).
80. Brenke, R. *et al.* Application of asymmetric statistical potentials to antibody–protein docking. *Bioinformatics* **28**, 2608–2614 (2012).
81. Desta, I. T., Porter, K. A., Xia, B., Kozakov, D. & Vajda, S. Performance and Its Limits in Rigid Body Protein-Protein Docking. *Structure* **28**, 1071–1081.e3 (2020).
82. Baek, M., Park, T., Heo, L., Park, C. & Seok, C. GalaxyHomomer: a web server for protein homooligomer structure prediction from a monomer sequence or structure. *Nucleic Acids Res* **45**, W320–W324 (2017).

83. Ko, J., Park, H., Heo, L. & Seok, C. GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Res* **40**, W294-297 (2012).
84. Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-Step Atomic-Level Energy Minimization: Biophysical Journal. [https://www.cell.com/biophysj/fulltext/S0006-3495\(11\)01245-8](https://www.cell.com/biophysj/fulltext/S0006-3495(11)01245-8).
85. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Research* **32**, W665–W667 (2004).
86. Ortega, A., Amorós, D. & García de la Torre, J. Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophysical Journal* **101**, 892–898 (2011).
87. Bertelsen, E. B., Chang, L., Gestwicki, J. E. & Zuiderweg, E. R. P. Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *PNAS* **106**, 8471–8476 (2009).
88. Bank, R. P. D. RCSB PDB - 4G6F: Crystal Structure of 10E8 Fab in Complex with an HIV-1 gp41 Peptide. <https://www.rcsb.org/structure/4G6F>.
89. Pinotsis, N. *et al.* Superhelical Architecture of the Myosin Filament-Linking Protein Myomesin with Unusual Elastic Properties. *PLOS Biology* **10**, e1001261 (2012).
90. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 1–11 (2021) doi:10.1038/s41586-021-03819-2.
91. AlphaFold Protein Structure Database. <https://www.alphafold.ebi.ac.uk/entry/P0A6Y8>.
92. Huang, J. *et al.* Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature* **491**, 406–412 (2012).
93. SIRAH forcefield in GROMACS. <http://www.sirahff.com/2012/09/sirah-forcefield-in-gromacs.html>.
94. Patel, J. S. & Ytreberg, F. M. Fast Calculation of Protein–Protein Binding Free Energies Using Umbrella Sampling with a Coarse-Grained Model. *J Chem Theory Comput* **14**, 991–997 (2018).
95. Durchschlag, H. Specific Volumes of Biological Macromolecules and Some Other Molecules of Biological Interest. in *Thermodynamic Data for Biochemistry and Biotechnology* (ed. Hinz, H.-J.) 45–128 (Springer, 1986). doi:10.1007/978-3-642-71114-5_3.
96. Bardwell, J. C. & Craig, E. A. Major heat shock gene of Drosophila and the Escherichia coli heat-inducible dnaK gene are homologous. *PNAS* **81**, 848–852 (1984).
97. Bolanos-Garcia, V. M. & Davies, O. R. Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1760**, 1304–1313 (2006).
98. Kozakov, D., Brenke, R., Comeau, S. R. & Vajda, S. PIPER: an FFT-based protein docking program with pairwise potentials. *Proteins* **65**, 392–406 (2006).
99. Černý, J. *Osobní konzultace*.
100. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery* **10**, 449–461 (2015).
101. Hou, T., Wang, J., Li, Y. & Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model* **51**, 69–82 (2011).

